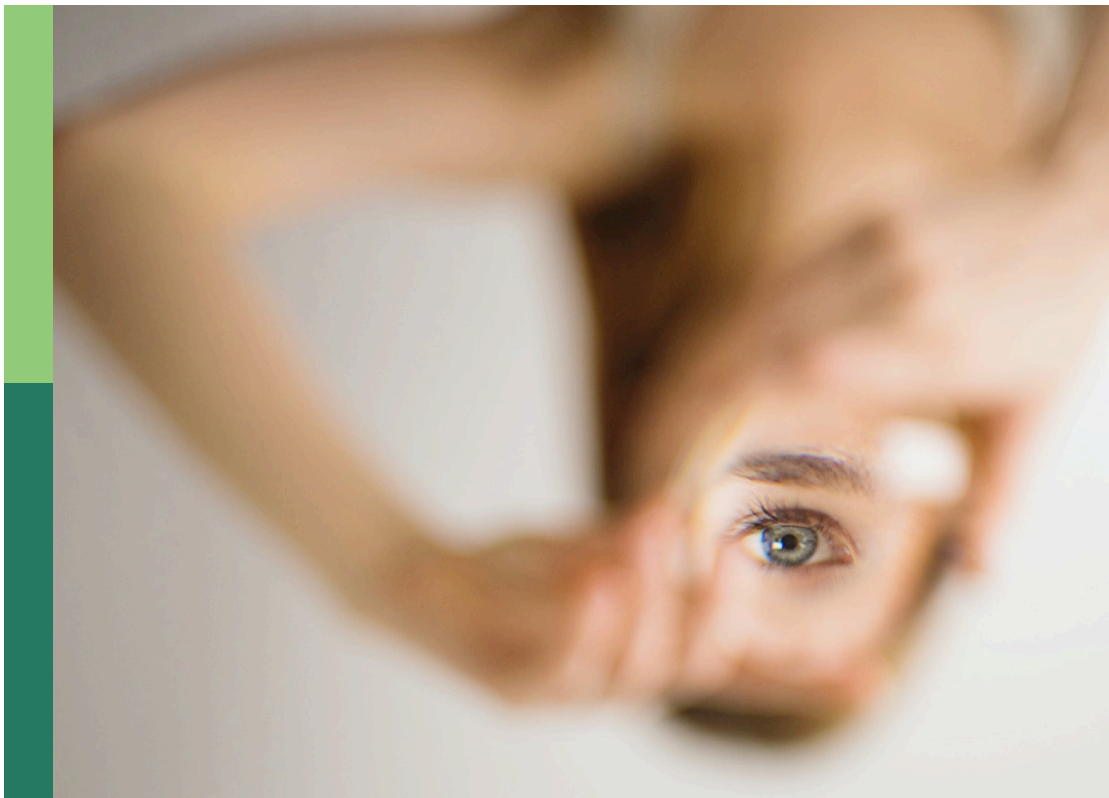
frontiers | Research Topics

Effective and attractive communication signals in social, cultural, and business contexts

Edited by
Oliver Niebuhr, Francesca D'Errico, Ellen Anke Schmid, Alexander Michael Brem and Anna Esposito

Published in
Frontiers in Psychology
Frontiers in Communication
Frontiers in Computer Science



FRONTIERS EBOOK COPYRIGHT STATEMENT

The copyright in the text of individual articles in this ebook is the property of their respective authors or their respective institutions or funders. The copyright in graphics and images within each article may be subject to copyright of other parties. In both cases this is subject to a license granted to Frontiers.

The compilation of articles constituting this ebook is the property of Frontiers.

Each article within this ebook, and the ebook itself, are published under the most recent version of the Creative Commons CC-BY licence. The version current at the date of publication of this ebook is CC-BY 4.0. If the CC-BY licence is updated, the licence granted by Frontiers is automatically updated to the new version.

When exercising any right under the CC-BY licence, Frontiers must be attributed as the original publisher of the article or ebook, as applicable.

Authors have the responsibility of ensuring that any graphics or other materials which are the property of others may be included in the CC-BY licence, but this should be checked before relying on the CC-BY licence to reproduce those materials. Any copyright notices relating to those materials must be complied with.

Copyright and source acknowledgement notices may not be removed and must be displayed in any copy, derivative work or partial copy which includes the elements in question.

All copyright, and all rights therein, are protected by national and international copyright laws. The above represents a summary only. For further information please read Frontiers' Conditions for Website Use and Copyright Statement, and the applicable CC-BY licence.

ISSN 1664-8714
ISBN 978-2-8325-2761-0
DOI 10.3389/978-2-8325-2761-0

About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

Effective and attractive communication signals in social, cultural, and business contexts

Topic editors

Oliver Niebuhr — University of Southern Denmark, Denmark
Francesca D'Errico — University of Bari Aldo Moro, Italy
Ellen Anke Schmid — Munich University of the Federal Armed Forces, Germany
Alexander Michael Brem — University of Stuttgart, Germany
Anna Esposito — University of Campania 'Luigi Vanvitelli, Italy

Citation

Niebuhr, O., D'Errico, F., Schmid, E. A., Brem, A. M., Esposito, A., eds. (2023). *Effective and attractive communication signals in social, cultural, and business contexts*. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-8325-2761-0

Topic Editor Oliver Niebuhr is the founder and CEO of AllGoodSpeakers ApS. The other Topic Editors declare no competing interests with regard to the Research Topic subject.

Table of contents

06 Editorial: Effective and attractive communication signals in social, cultural, and business contexts
Oliver Niebuhr, Francesca D'Errico, Anna Esposito, Ellen Schmid and Alexander Brem

11 Emotivity in the Voice: Prosodic, Lexical, and Cultural Appraisal of Complaining Speech
Maël Mauchand and Marc D. Pell

24 Effects of Self-Expressive Brand and Susceptibility to Interpersonal Influence on Brand Addiction: Mediating Role of Brand Passion
Shizhen Bai, Yue Yin, Yubing Yu, Sheng Wei and Rong Wu

35 Sound-Symbolic Semantics of Written Japanese Vowels in a Paper-Based Survey Study
Misa Ando, Xinyi Liu, Yan Yan, Yutao Yang, Shushi Namba, Kazuaki Abe and Toshimune Kambara

44 The Impacts of Perceived Risk and Negative Emotions on the Service Recovery Effect for Online Travel Agencies: The Moderating Role of Corporate Reputation
Jiahua Wei

58 Chinese App User's Needs Profile: From Questionnaire Measurement to Behavior Analysis
Zaoyi Sun, Liang Xu, Qi Zhong and Xiuying Qian

69 Cyber-Personality and *Liking* Expression: A Study From WeChat Users in China
Haojian Li and X.T. Wang

78 Using E-GARCH to Analyze the Impact of Investor Sentiment on Stock Returns Near Stock Market Crashes
Sze Ting Chen and Kai Yin Allison Haga

94 Emotional Gaze: The Effects of Gaze Direction on the Perception of Facial Emotions
Jing Liang, Yu-Qing Zou, Si-Yi Liang, Yu-Wei Wu and Wen-Jing Yan

100 The Humble Charisma of a White-Dressed Man in a Desert Place: Pope Francis' Communicative Style in the Covid-19 Pandemic
Rosa Scardigno, Concetta Papapicco, Valentina Luccarelli, Altomare Enza Zagaria, Giuseppe Mininni and Francesca D'Errico

108 Embedded Philanthropic CSR in Digital China: Unified View of Prosocial and Pro-environmental Practices
Qing Ye, Zain Rafique, Rongting Zhou, Fahad Asmi, Muhammad Azfar Anwar and Ahmad Nabeel Siddiquei

121 Shape–Trait Consistency: The Matching Effect of Consumer Power State and Shape Preference
Feng Yao, Xiaotong Jin, Banggang Wu, Taiyang Zhao and Tiannv Ma

132 Hong Kong Women Project a Larger Body When Speaking to Attractive Men
Albert Lee and Eva Ng

139 A Study on the Correlation Between Media Usage Frequency and Audiences' Risk Perception, Emotion and Behavior
Peng-Peng Li and Fangqi Zhong

154 Effects of Social Connectedness on the Sharing of Employee-Created Content
Xueting Zhang and Man Chen

167 Coordinating Mechanisms Are More Important Than Team Processes for Geographically Dispersed Emergency Dispatch and Paramedic Teams
Bjørn Helge Johnsen, Roar Espevik, Jarle Eid, Øyvind Østerås, Johan Kolstad Jacobsen and Guttorm Brattebø

177 Prosodic Phrasing of Good Speakers in English and Czech
Radek Skarnitzl and Hana Hledíková

190 Research on Bond Participants' Emotion Reactions Toward the Internet News in China's Bond Market
Wei Zhang, Jun Wang and Mu Tong

203 Pleasantness and Wellbeing in Poem Declamation in European and Brazilian Portuguese Depends Mostly on Pausing and Voice Quality
Plinio A. Barbosa

216 Effects of Social Media Usage on Consumers' Purchase Intention in Social Commerce: A Cross-Cultural Empirical Analysis
Shangui Hu and Zhen Zhu

233 The Effects of Leader Emotional Intelligence, Leadership Styles, Organizational Commitment, and Trust on Job Performance in the Real Estate Brokerage Industry
Chun-chang Lee, Yei-shian Li, Wen-chih Yeh and Zheng Yu

254 Product-Centered Language Style in Chinese Women Online Reviews
Xian Wang, Congjun Mu, Huixian Li, Alison Noble and Qingyi Wang

264 Effectiveness, Attractiveness, and Emotional Response to Voice Pitch and Hand Gestures in Public Speaking
Emma Rodero

276 Research on the Visual Imagery of Posters Based on the Culture Code Theory of Design
Luo Wen, Wang Jingjing, Wang Chen and Sun Luyu

282 Multimodal Prominence Marking in Semi-Spontaneous YouTube Monologs: The Interaction of Intonation and Eyebrow Movements
Stephanie Berger and Margaret Zellers

301 President Obama's Humble Face: An Authentic or a Socially Desirable Posturing? A Study on Reactions to Obama's Autobiographical Self-Disclosures
Alessia Mastropietro, Peter Bull, Francesca D'Errico, Isora Sessa, Stefano Migliorisi and Giovanna Leone

308 Laugh is in the air: An exploratory analysis of laughter during speed dating
Tamara Rathcke and Susanne Fuchs

324 Offensive or amusing? The study on the influence of brand-to-brand teasing on consumer engagement behavioral intention based on social media
Yu-mei Ning, Chuan Hu, Ting-ting Tu and Dan Li

341 How to motivate with speech. Findings from acoustic phonetics and pragmatics
Jana Voße, Oliver Niebuhr and Petra Wagner

349 Thoughts on the usage of audible smiling in speech synthesis applications
Jürgen Trouvain and Benjamin Weiss

356 Unguided virtual-reality training can enhance the oral presentation skills of high-school students
Ïo Valls-Ratés, Oliver Niebuhr and Pilar Prieto

377 Is wearing high heels a female mating strategy? Revisiting the original study using qualitative methods
Radomír Masaryk, Nikola Synak and Michaela Belišová

385 Cross-platform opinion dynamics in competitive travel advertising: A coupled networks' insight
Jia Chen, Haomin Wang and Xiangrui Chao

396 Evolutionary game analysis of FinTech transformation: A social co-governance pattern of peer-to-peer lending market in China
Qi Wang, Xin Liu and Chenghu Zhang

420 The effects of different voice qualities on the perceived personality of a speaker
Sara Pearsell and Daniel Pape

440 A digital "flat affect"? Popular speech compression codecs and their effects on emotional prosody
Oliver Niebuhr and Ingo Siegert

frontiers | Frontiers in Psychology

TYPE Editorial
PUBLISHED 16 May 2023
DOI 10.3389/fpsyg.2023.1205329

Check for updates

OPEN ACCESS

EDITED AND REVIEWED BY
Xiaolin Zhou,
Peking University, China

*CORRESPONDENCE
Oliver Niebuhr
✉ olni@sdu.dk

RECEIVED 13 April 2023
ACCEPTED 18 April 2023
PUBLISHED 16 May 2023

CITATION
Niebuhr O, D'Errico F, Esposito A, Schmid E and Brem A (2023) Editorial: Effective and attractive communication signals in social, cultural, and business contexts. *Front. Psychol.* 14:1205329. doi: 10.3389/fpsyg.2023.1205329

COPYRIGHT
© 2023 Niebuhr, D'Errico, Esposito, Schmid and Brem. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Editorial: Effective and attractive communication signals in social, cultural, and business contexts

Oliver Niebuhr[1]*, Francesca D'Errico[2], Anna Esposito[3], Ellen Schmid[4] and Alexander Brem[5,6]

[1]Centre for Industrial Electronics, University of Southern Denmark, Sonderborg, Denmark, [2]Department of Educational Sciences, Psychology and Communication, University of Bari 'Aldo Moro', Bari, Italy, [3]International Institute for Advanced Scientific Studies (IIASS), Department of Psychology, Università della Campania "Luigi Vanvitelli", Naples, Italy, [4]Department of Business Administration, Universität der Bundeswehr München, Neubiberg, Germany, [5]University of Stuttgart, Institute of Entrepreneurship and Innovation Science (ENI), Stuttgart, Germany, [6]Department of Technology and Innovation, University of Southern Denmark, Sonderborg, Denmark

KEYWORDS

charisma, speech, gestures, leadership, branding (marketing), E-commerce, emotion, social media

Editorial on the Research Topic
Effective and attractive communication signals in social, cultural, and business contexts

According to current social media statistics, the number of YouTube users, including influencers, has doubled worldwide since 2017 (Gandola, 2022). During the period of our Research Topic alone, about 300 million people joined YouTube to consume and produce content. The number of podcasts and their listeners is growing constantly by around 20% for years now as well, prompting Richter (2022) to ask whether podcasts could become a new mass medium. A recent survey among executives of leading companies in 10 Western countries found 10 years ago already that almost 80% of them support having a business-related social media presence (Gesenhues, 2013). In fact, more than 70% of CEOs, like those from the German DAX-40 companies, regularly post on social media, especially LinkedIn, and use it more and more strategically, in this way becoming "corporate influencers" (Warschun and Zehnder, 2020). Finally, voice assistants and related talking machines are on their way to conquer our everyday lives. Since the start of our Research Topic, their number has almost doubled to an estimated 8.4 billion units worldwide (Laricchia, 2022). It is impressive statistics like these that motivated us to create the Frontiers Research Topic (RT) on effective and attractive communication signals.

The common idea behind our initiative was that communication signals are not only (and perhaps not even primarily) about exchanging propositions. We talk and text and listen and read not least for social reasons. Communication is essentially a social instrument; we seek attention, we define and change social hierarchies, we create and develop social bonds, and in all these contexts we share feelings, emotions, attitudes, tears, and laughter, even with voice assistants (D'Errico and Poggi, 2016; Poggi and D'Errico, 2022). Without the essentially social nature of communication, there might be no "small talk", storytelling would not be a successful rhetorical strategy, AI's would not be able to predict the success of romantic relationships or business meetings from linguistic and phonetic features, attempts to change a language (like in the debate on gender-neutral language) would take place without sparking emotionally charged debates, and commercials, poetry, and news would probably all look and sound identical.

The decisive point of our RT in this context was that the individual scientific disciplines dealing with effective and attractive communication signals have so far researched and

published largely independently of one another. Our RT therefore not only set itself the goal to advance the understanding of these signals and their effects, but also to make such original findings visible across disciplines and accessible in the form of open-access papers, in order to stimulate interactions and collaborations across disciplines. I think we have done remarkably well in achieving both of these goals!

I would like to thank our 111 authors for working so hard on their papers, and my active team of co-editors from the fields of Social Psychology (Francesca D'Errico), Behavioral Cognitive Systems (Anna Esposito), Entrepreneurship (Alexander M. Brem) and Leadership (Ellen A. Schmid) for guiding our authors so quickly and smoothly through their revision processes. Fifteen accepted papers were our original goal when the journey began in 2021. We ended the journey with 35 accepted papers—out of 55 submissions. The rejection rate of almost 40% speaks for the conscientious work of the many reviewers, to whom we are also very grateful, and it reflects the high quality and innovativeness of the remaining contributions—most of which have already achieved an above-average number of views by Frontier's standards.

In terms of content, the effective and attractive communication signals investigated in this RT range from the language of social commerce and brand passion (Bai et al.) to strategically applicable sound-meaning associations of Japanese vowels (Ando et al.) to women's intention behind wearing high heels (Masaryk et al.).

Emotions in particular, run like a red thread through the chronology and disciplines of the RT. More than half of all contributions to the RT are directly or indirectly related to the production or perception of human emotions. Let us follow this thread to illustrate how diversely and substantially the RT has advanced our understanding of effective and attractive communication signals. Beforehand, however, I would like to ask all authors for their understanding that, given the word limit of this editorial, we cannot portray all papers equally. It does not mean that we value the papers less that analyzed connections between, for example, intonation and eyebrow movements in YouTubers' realization of word stress (Berger and Zellers), between cyber-personality and liking expression (Li and Wang), between voice acoustics and motivational speech (Voße et al.), or between social media usage and consumers' purchase intention (Hu and Zhu). The conclusions that we draw from the emotion studies portrayed below actually apply in similar ways to all these other papers as well.

Liang et al. show, for example, that gaze direction significantly influences the perception of emotions. As stimuli, they used photographs of people either looking directly into the participants' eyes or avoiding eye contact with the participants. People with direct eye contact were more likely to be attributed positive emotions by participants, while avoiding eye contact was associated with negative emotions of fear or anger—notably, this was true for both smiling and neutral faces in the photos. These results are not fully consistent with previous studies. Liang et al. attribute that to the Asian culture where the experiment took place (as compared to previous experiments conducted in Western contexts). Mauchand and Pell did not examine images, but sounds, more precisely, the tone-of-voice of speakers who aim to trigger empathy in interlocutors through third-party complaints. In addition to the difference between a neutral vs. an emotional (complaining) tone of voice, the verbal content of the stimuli and the social status of the speakers were varied as well. Mauchand and Pell found only small effect sizes for influences of social status and verbal content on the interlocutors' empathy responses. That is, "what characterizes a complaint is how it is said [...], more than what it is about or who produces it". The two researchers thus draw conclusions similar to those drawn in connection with charismatic speech (cf. Caspi et al., 2019) and even hate speech (Paciello et al., 2021; cf. Niebuhr and Neitsch, 2022).

Voice patterns, i.e. prosodies, also played a role in several other RT contributions. Pearsell and Pape analyzed the relationships between prosody and the Big-5 personality traits, specifically those traits associated with perceived speaker charisma. Unlike earlier studies, they focused on the 4th dimension of prosody, viz. timbre (see Campbell and Mokhtari, 2003). In agreement with previous studies, Pearell and Pape found that a smiling voice is positive for perceived charisma, while a creaky voice significantly undermines the speaker's perceived charisma (see also Pointer et al., 2022; Tschinse et al., 2022). Both effects were more pronounced for female than male voices, at least in combination with short single-statement audio stimuli; for longer audio stimuli this might no longer apply, see Tschinse et al. (2022).

The contribution by Trouvain and Weiss to the RT elaborates on smiling, which is becoming a hot topic in the speech sciences. Trouvain and Weiss discuss, with special emphasis on synthetic voices, how (auditory) smiles can improve audio books, social robots, and dialogue systems in the future, and what acoustic requirements have to be met to that end. Trouvain and Weiss concluded their paper with a call-to-action: before we can properly implement smiling into synthetic speech "basic research is needed with respect to (i) when exactly, (ii) to which degree, and (iii) for which purpose humans smile in spoken interaction". In the paper by Niebuhr and Siegert, voice and technology are not viewed from a constructive but from a destructive perspective. With reference to previous perception results, the two researchers found that speakers who make cell-phone or video calls at high speech-compression rates can be subject to "digital flat affect". That is, listeners are no longer able to recognize the emotions that those speakers intended to convey. This even applies to strong basic emotions such as happiness, anger, fear, or boredom. In their RT contribution, Niebuhr and Siegert identified the acoustic distortions underlying this "digital flat affect". Results showed that the acoustic distortions of intonation were less responsible for the "digital flat effect" than those of timbre, which suggests that emotions are acoustically encoded in timbre patterns to an extent that has perhaps been underestimated so far, similar to what Pearsell and Pape suspect as well (see above).

Four other papers look at the voice-emotion link from a more performance-oriented angle. Based on the assumption that TED talkers are "good speakers", Skarnitzl and Hledíková compared the vocal performance indicators of Czech and English TED talks, 10 per language. In an acoustic analysis, they found that English TED talkers had more melodic variability and paused more often than Czech ones. Skarnitzl and Hledíková discuss these findings with reference to a culture and language-specific perception of perceived speaker charisma (see Biadsy et al., 2008). Barbosa examined how the acoustics of recited poetry correlates with the perception of

pleasantness and wellbeing in listeners. To that end, he made Brazilian and European Portuguese speakers read the same poem. An extensive acoustic analysis showed that it was primarily pausing (pause duration, pausing rate) and timbre (spectral emphasis and LTAS slope) that correlated with the listener ratings. The voice pitch level also played a role. Specifically, at least for poetry reading, a feeling of pleasantness and wellbeing was best conveyed by very low (or very high) and overall softer voices. Speaking slowly with long pauses also contributed to triggering pleasantness and wellbeing. By contrast, the study of Valls-Ratés et al. was about the exact opposite type of feelings: fear and stress or, more precisely, people's very widespread public-speaking anxiety (PSA). Based on a large-scale study that combined acoustic analyses and ratings, the researchers showed that, on the one hand, VR-based public-speaking training was able to trigger a PSA response similar to that of a real audience; on the other hand, only a few minutes of VR-based public-speaking training per week were already sufficient to significantly reduce the PSA level of speakers and, unlike a non-VR control group, the VR speakers also considered their training beneficial for their own performance. This evaluation was consistent with the acoustic analysis of Valls-Ratés et al., according to which the VR speakers' voices became more sonorous and melodic and their speaking rates slower and calmer after training.

From reduced anxiety, it is not far along our red thread until attractiveness and fun. Studies on such communication signals are also included in our RT. For example, Rathcke and Fuchs examined laughter in a heterosexual speed-dating setting. To that end, they arranged and recorded 12 dates between single men and women, who, in addition, also took part in same-sex conversations as a baseline condition. After each session, participants rated the interlocutor's romantic attractiveness. Rathcke and Fuchs concluded from their data that romantic attraction is indeed indicated to a certain degree by laughter and can thus be predicted by laughter. Specifically, "(a) laughs are particularly frequent in the first minute of the conversation, (b) daters who are mutually attracted show a significantly larger degree of temporal overlap in laughs, (c) specific laughter types (classified as a nasal 'laugh-snort') prevail in high-attraction dates". Lee and Ng also analyzed romantic attractiveness, albeit for Cantonese rather than German—and moreover, not in terms of laughter, but in terms of acoustically projected body size. Female participants rated photos of male faces for visual attractiveness and then produced short utterance sequences related to these faces. Results showed that, a higher visual attractiveness of male faces made female participants speak with less breathy, lower-pitched voices and such that the formant (i.e. resonance) frequencies in their speech were lower and closer together. That is, they spoke with an overall voice pattern suitable for conveying a larger acoustic body size. These effects were stronger the more satisfied the female speakers were with their own height. Finally, Rodero's study reiterates the emotions associated with perceived speaker charisma and additionally, comes full circle back to the study by Liang et al. as it includes visual behavioral patterns. Rodero made speakers combine different hand/arm gesture amplitudes with different degrees of melodic variation and let 120 participants rate the resulting A/V stimuli, both explicitly via rating scales and implicitly by measuring the participants' skin-conductance response (EDA). Rodero's results were consistent with Rosenberg and Hirschberg (2009), who have already pointed out the potential risk of exaggerating non-verbal parameters in the perception of speaker charisma, and Michalsky and Niebuhr (2019), who have experimentally defined these overdose thresholds for numerous prosodic parameters. Rodero also found that, perceptually, the most effective and attractive speakers were characterized by moderate levels of gesturing and melodic variability. Speakers with strongly pronounced gestures performed best in terms of EDA biosignals, but only when their melodic variability was low or moderate. In any case, speaking with strongly pronounced gestures and with a high level of melodic variability was perceived as neither effective nor attractive.

Powerful public speaking is undoubtedly closely related to commercial activities, and our RT also features emotion-related papers on such activities. For example, supporting the findings of previous studies but for a different business sector and cultural environment, Lee et al. report the following: Transformational, i.e. essentially charismatic business leaders, can positively influence their employees' job performance, and more generally, the emotional intelligence of business leaders is positively correlated with employees' trust in them and, interestingly, also in each other. Employees' trust in each other furthermore determines the extent to which the business leaders' emotional intelligence can positively affect employees' job performance. From the passion of leaders and employees, we proceed to Bai et al. and the passion for brands and how intense, highly emotional consumer–brand relationships can be established. Bai et al. showed how self-expressiveness and susceptibility to interpersonal influence contribute to creating an intense consumer–brand relationship. Specifically, based on their data, they stress that brand "managers should consider customers' desire for self-expression and their need to connect with others" and that having a "passion branding strategy" is essential for the success of a product, at least in the investigated Asian context. A particularly emotionally charged type of communication signal is humor, especially black humor. Ning et al. analyzed the influence of brand-to-brand teasing on consumer engagement, by means of systematically varied Twitter tweets in a large-scale perception experiment. They found that low-aggressive humor can increase consumer engagement for one's own brand, especially if the humor is presented with a wink and matches with the company's own brand personality.

Let us move on from good and bad humor to good and bad news. Zhang et al. studied how credit rating agencies on the one hand, and bond investors on the other, deal with good and bad news. From a large-scale field study for which almost 250,000 news articles were collected and analyzed over 10 years (2010-2020), Zhang et al. conclude that credit rating agencies and bond investors behave fundamentally differently, at least in China. While bond investors react very sensitively to bad news, credit rating agencies largely ignore such news; and, compared to credit rating agencies, bond investors react more strongly to bad than to good news, and are generally less sensitive to news from state-owned companies. All about bad news is Wei's study on failed bookings with online travel agencies. Wei empirical research shows that the higher the level of negative emotions caused by failed bookings, and the worse the booking agency reputation, the weaker the service recovery

satisfaction is for customers. In practical terms, this means that online travel agencies must act quickly after a failed booking in order to prevent negative emotions building up in customers, and they must increase their reputation in order to have a loyalty buffer with customers. Media also plays a role in the study by Li and Zhong. Based on field data, the two researchers evaluate the media's influence on risk perception, emotions and behavior of the Chinese population during the COVID-19 pandemic. Their results showed, in the form of correlation analyses, "that the degree (frequency) of people's use of different media in…COVID-19…significantly affected their negative emotions of fear, worry, and anxiety." The data showed that negative emotions were more readily influenced by online media as compared to "interpersonal and television [influences], while positive emotions were mainly influenced by people's use of television".

The study by Scardigno et al. was motivated by the COVID-19 pandemic as well. The aim was "to shed light on the communicative features of the charismatic leadership of Pope Francis during the pandemic emergency" by means of analyzing both his discursive rhetorical figures and "his multimodal body signals in the global TV event of the Universal Prayer with the Urbi et Orbi Blessing". Scardigno et al. reveal and define Pope Francis' "humble charisma". He combines open posture and inclusive wording (e.g. the use of "we") with emotionally charged language and common rhetorical figures such as metaphors and tri-part lists on the one hand, and with expressions of weakness and humility on the other (D'Errico and Poggi, 2019). Finally, the last study portrayed here by Mastropietro et al. is also about humility; or rather, authenticity, because what the researchers showed in their cross-modal perception experiment with 175 participants was that Obama's humble charisma (determined in previous studies) was only effective and rated as authentic if his facial expression—varied in terms of photos—was emotionally consistent with the content of speech excerpts read by the participants. In the case of mismatches, Obama's facial expression was considered "simple socially desirable posturing".

Given all these RT studies that revolve around the production and perception of emotions, what fundamental conclusions can we draw about the nature of the communication signals analyzed? Clearly, the RT underpinned that the communication signals we all send out are (beyond their mere information content), always more or less effective and attractive. Communication signals, be they written or spoken, can determine what we buy, who we like, who we trust and forgive, whether we feel or overcome fear, and who we pay attention to commercially or socially. The signals are multimodal and to some extent both similar and different across cultures and languages. They can be positive and negative, with negative signals not necessarily evoking negative or unwanted reactions. For instance, speakers can come across as charismatic even if they convey righteous anger. In addition, culture and language-specific overdose thresholds exist and must be defined and taken into account. Related to this, it is apparent that effective and attractive communication signals must be internally consistent in order to unfold their (intended) effect on recipients. All of the above implies that effective and attractive communication signals are worth being trained and that they need to be well mastered. This applies not only in social contexts, but perhaps even more in economic and political contexts.

Looking ahead, given the multi-channel nature and contextual complexity of effective and attractive signals, such good mastery seems a long way off, both in theory and in practice, and particularly for voice assistants and other talking machines. As was stated by Trouvain and Weiss, we first have to better understand how these signals are used among us humans before we can—context-sensitively—identify patterns and transfer them to machines (Castellano et al., 2021). There is still a lot to learn and to research, especially regarding the links between communication channels (text, speech, "body language") and the timbre of voice acoustics (which includes laughter and smiles). The same applies to the digital technology that we create and use to transmit our effective and attractive communication signals and to support our training of these signals, for example by means of VR.

My co-editors Francesca D'Errico, Anna Esposito, Ellen A. Schmid, Alexander M. Brem, and I hope that this RT can make a lasting contribution to connecting the relevant scientific disciplines with regard to effective and attractive communication signals, both in terms of their exploration as well as in terms of their quantification, evaluation, and training.

The English satirist Charlton Brooker once said that, "in the age of social media, everyone's a newspaper columnist". In a similar context, the US American author and speaker pointed out that "we aren't in an information age, we are in an entertainment age". It is up to us researchers, designers, and developers of effective and attractive communication signals to live up to the implicit responsibility expressed in these quotes. I think we are—together—on an effective and attractive way to achieve this goal.

Author contributions

ON authored the editorial on behalf of all co-editors. ON, FD'E, AE, ES, and AB contributed equally to the conception, design, and organization of the Research Topic. ON, FD'E, AE, and ES additionally edited and/or reviewed manuscripts for the Research Topic. ON and FD'E are furthermore co-authors of manuscripts published in this Research Topic. All authors contributed to the article and approved the submitted version.

Acknowledgments

My co-editors and I are very grateful to Charlotte Fooks for proof-reading an earlier draft of this editorial. Moreover, thanks are due to the editorial team of Frontiers for all the quick, effective, and understanding correspondence in connection with this Research Topic.

Conflict of interest

ON is the CEO of the speech technology company AllGoodSpeakers APS, please find further details here: https://oliverniebuhr.com/conflict-of-interst.html.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Biadsy, F., Rosenberg, A., Carlson, R., Hirschberg, J., and Strangert, E. (2008). A cross-cultural comparison of American, Palestinian, and Swedish perception of charismatic speech. *Proc. Int. Conf. Speech Prosody* 37, 579–582.

Campbell, N., and Mokhtari, P. (2003). Voice quality: the 4th prosodic dimension. *Proc. Int. Cong. Phonetic Sci.* 15, 2417–2420.

Caspi, A., Bogler, R., and Tzuman, O. (2019). "Judging a book by its cover": the dominance of delivery over content when perceiving charisma. *Group Org. Manage.* 44, 1067–1098. doi: 10.1177/1059601119835982

Castellano, G., Carolis, D., e., D'Errico, B., and Macchiarulo, F. N., and Rossano, V. (2021). PeppeRecycle: improving children's attitude toward recycling by playing with a social robot. *Int. J. Soc. Robotics* 13, 97–111. doi: 10.1007/s12369-021-00754-0

D'Errico, F., and Poggi, I. (2016). "The bitter laughter". When parody is a moral and affective priming in political persuasion. *Front. Psychol.* 7, 1144. doi: 10.3389/fpsyg.2016.01144

D'Errico, F., and Poggi, I. (2019). *Tracking a Leader's Humility and Its Emotions From Body, Face and Voice. Web Intelligence*. London: IOS Press.

Gandola, K. (2022). *How Well Does YouTube Influencer Marketing Work? Net Influencer*. Available online at: https://www.netinfluencer.com/youtube-influencer-marketing/ (accessed 7 April, 2023).

Gesenhues, A. (2013). *Study: 76% Of Executives Believe CEOs Should Be Utilizing Social Media Channels. Martech, Data Topics*. Available online at: https://martech.org/study-ceos-active-on-social-media-channels-rate-higher-as-leaders-compared-to-ceos-not-using-social-media/ (accessed 7 April, 2023).

Laricchia, F. (2022). *Number of Digital Voice Assistants in Use Worldwide 2019-2024. Statistica – Consumer Electronics*. Available online at: https://www.statista.com/statistics/973815/worldwide-digital-voice-assistant-in-use/ (accessed 7 April, 2023).

Michalsky, J., and Niebuhr, O. (2019). Myth busted? Challenging what we think we know about charismatic speech. *AUC Philolo.* 2019, 27–56. doi: 10.14712/24646830.2019.17

Niebuhr, O., and Neitsch, J. (2022). The truth below the surface: Towards quantifying and understanding the evaluation of German and Danish hate speech with EEG biosignals. *J. Speech Sci.* 11, e022004–e022004. doi: 10.20396/joss.v11i00.16153

Paciello, M., D'Errico, F., Saleri, G., and Lamponi, E. (2021). Online sexist meme and its effects on moral and emotional processes in social media. *Comput. Hum. Behav.* 116, 106655. doi: 10.1016/j.chb.2020.106655

Poggi, I., and D'Errico, F. (2022). *Social Influence, Power, and Multimodal Communication*. New Yor, NY: Taylor and Francis.

Pointer, N. F., van Mersbergen, M., and Nanjundeswaran, C. D. (2022). *Listeners' attitudes towards young women with glottal fry. J. Voice.* doi: 10.1016/j.jvoice.2022.09.007

Richter, F. (2022). *Podcasts: The New Mass Medium? Statistica Topics*. Available online at: https://www.statista.com/chart/14306/podcast-adoption/ (accessed 7 April 2023).

Rosenberg, A., and Hirschberg, J. (2009). Charisma perception from text and speech. *Speech Communication* 51, 640–655. doi: 10.1016/j.specom.2008.11.001

Tschinse, L. M., Asadi, A., Gutnyk, A., and Niebuhr, O. (2022). Keep on smiling...? An exploration of the gender-specific connections between smiling duration and perceived speaker attributes in business pitches. *Proc. 11th Int. Conf.* 22, 624–628. doi: 10.21437/SpeechProsody.2022-127

Warschun, M., and Zehnder, E. (2020). *CXO als Corporate Influencer – Digital Branding Studie 2020*. Available online at: https://www.egonzehnder.com/cdn/serve/article-pdf/1606922771-538a8fb0e2a03f916566d086944fe810.pdf (accessed 7 April, 2023).

frontiers in Psychology

ORIGINAL RESEARCH
published: 18 January 2021
doi: 10.3389/fpsyg.2020.619222

Emotivity in the Voice: Prosodic, Lexical, and Cultural Appraisal of Complaining Speech

Maël Mauchand and Marc D. Pell*

School of Communication Sciences and Disorders, McGill University, Montreal, QC, Canada

Emotive speech is a social act in which a speaker displays emotional signals with a specific intention; in the case of third-party complaints, this intention is to elicit empathy in the listener. The present study assessed how the emotivity of complaints was perceived in various conditions. Participants listened to short statements describing painful or neutral situations, spoken with a complaining or neutral prosody, and evaluated how complaining the speaker sounded. In addition to manipulating features of the message, social-affiliative factors which could influence complaint perception were varied by adopting a cross-cultural design: participants were either Québécois (French Canadian) or French and listened to utterances expressed by both cultural groups. The presence of a complaining tone of voice had the largest effect on participant evaluations, while the nature of statements had a significant, but smaller influence. Marginal effects of culture on explicit evaluation of complaints were found. A multiple mediation analysis suggested that mean fundamental frequency was the main prosodic signal that participants relied on to detect complaints, though most of the prosody effect could not be linearly explained by acoustic parameters. These results highlight a tacit agreement between speaker and listener: what characterizes a complaint is how it is said (i.e., the tone of voice), more than what it is about or who produces it. More generally, the study emphasizes the central importance of prosody in expressive speech acts such as complaints, which are designed to strengthen social bonds and supportive responses in interactive behavior. This intentional and interpersonal aspect in the communication of emotions needs to be further considered in research on affect and communication.

Keywords: pragmatics, cross-cultural, vocal affect, emotive involvement, complaint

OPEN ACCESS

Edited by:
Oliver Niebuhr, University of Southern Denmark, Denmark

Reviewed by:
Albert Rilliard, UPR3251 Laboratoire d'Informatique pour la Mécanique et les Sciences de l'Ingénieur (LIMSI), France
Joan Ma, Queen Margaret University, United Kingdom

****Correspondence:***
Maël Mauchand
mael.mauchand@mail.mcgill.ca

Specialty section:
This article was submitted to Language Sciences, a section of the journal Frontiers in Psychology

Received: *19 October 2020*
Accepted: *21 December 2020*
Published: *18 January 2021*

Citation:
Mauchand M and Pell MD (2021) Emotivity in the Voice: Prosodic, Lexical, and Cultural Appraisal of Complaining Speech. Front. Psychol. 11:619222. doi: 10.3389/fpsyg.2020.619222

INTRODUCTION

While the role of the voice in social interaction has been receiving growing interest over the last decades, literature on the topic has been scattered across research fields. On the one hand, experimental psychology has been focusing on affective speech and emotions (Frick, 1985; Scherer, 2003; Juslin, 2013); on the other hand, intentionality, speech acts, and attitudes have been mostly addressed by pragmatics and theoretical linguistics (Searle, 1965; Grice, 1989; Wichmann, 2002; Culpeper and Terkourafi, 2017). A large part of our daily social interactions is inherently *emotive*, relying on the attitudinal, intentional use of emotional signals (Caffi and Janney, 1994).

These interactions, involving both speaker and listener in a complex collaborative timeline, remain poorly understood. The nature and components of emotive interactions can be investigated through an intersectional approach, embedding social and affective psychology methods into the theoretical pragmatics framework of emotivity through the Emotions As Social Information (EASI) model (Van Kleef, 2009). Focusing on the case of complaints, the present study examines how emotivity is conveyed through speech, and how affective signals in the voice are processed in different social and cultural contexts.

Complaints are intentional verbal expressions of social pain, distress, or displeasure (Boxer, 1993; Drew, 1998; Laforest, 2002), and are usually divided in two categories. *Direct* complaints are addressed directly to the source of the issue, with the purpose of terminating or solving the issue (Trenchs, 1994; Laforest, 2002). The present study focuses on *indirect* or *third-party* complaints, which are addressed to a third party usually unrelated to the issue (Drew, 1998; Edwards, 2005). Third-party complaints are non-instrumental in nature (Alicke et al., 1992); they do not aim to solve the problem they address, but have a more indirect function of promoting social affiliation though affectivity and empathy (Drew and Walker, 2009; Ogden, 2010). In what follows, the term complaint will be used to refer exclusively to third-party complaints.

The social importance of complaints is implied by their frequency; it is said that individuals complain more than four times a day on average (Alicke et al., 1992). While many types of speech acts can lead to a strengthening of social bonds, complaints appear to directly serve this purpose; this is accomplished through long, interactive sequences in which the complainer negotiates the affiliation of their listener (Drew and Walker, 2009; Selting, 2010). Complaints are usually defined by tightly bounded topics with a clear beginning and end (Drew, 1998), often used as ice-breakers or conversation openers (Boxer, 1993; Kowalski, 2002). Complainers may open with an initial complaint to probe the affiliative response of their listener, which will determine the course of the negotiation (Traverso, 2009). Ultimately, it is the listener who chooses whether or not to collaborate and affiliate with the speaker (Edwards, 2005; Selting, 2010).

Beyond describing a negative situation, a core function of complaints is to provide evidence of how the speaker *feels* about the situation (Drew, 1998). In order to gain affiliation, a complaint should allow the listener to share the affective state of the speaker and empathize with them (Acuña-Ferreira, 2002; Edwards, 2005). Since most complaints describe a past event involving felt pain or distress, it is unlikely that the speaker is fully experiencing these emotions as they complain; rather, these expressions may be viewed as instances of "reconstructed affect" (Selting, 2010). Complaining is thus an *emotive* or *expressive* speech act (Scarantino, 2017), a type of social performance in which the speaker intentionally displays affective markers to achieve interactive goals (Caffi and Janney, 1994; Acuña-Ferreira, 2002). These markers are the negotiating products of a complaint, informing the listener of the complainer's emotions (Edwards, 2005) and sharing these emotions through mood contagion (Kowalski, 2002). The affective component of a complaint is usually more important than the object of the complaint itself, from which the interaction can drift off while remaining a complaining collaboration (Edwards, 2005; Traverso, 2009).

The Emotions as Social Information (EASI) model (Van Kleef, 2009) provides a useful framework for investigating the perceptual and social dimensions of complaining speech in greater depth. The EASI model emphasizes that affective displays are more than biological symptoms and can be used to influence others by triggering inferential processes and affective reactions. For complaints to succeed (i.e., promote social affiliation and strengthen bonds), complainers and listeners need to effectively display, perceive and respond to communicative signals of affect and emotivity, which are frequently marked through a complainer's voice, or *speech prosody*. Here, we refer to prosody as suprasegmental acoustic features of speech - pitch, loudness, voice quality, rhythm - that speakers modulate, intentionally or not, to express meanings, emotions, and attitudes in their voice (Pell, 2001; Scherer, 2003). The manner in which prosody is used in complaining interactions and its impact on listeners has seldom been explored.

According to Brunswikian lens models of speech, the emotions of speakers are encoded by a constellation of acoustic cues that are then decoded by listeners into emotional representations (Brunswik, 1956; Grandjean et al., 2006; Laukka et al., 2016). A number of studies have reported that vocal expressions of basic emotion (e.g., anger, sadness, happiness) show specific patterns of pitch, loudness, rhythm, and voice quality that yield successful recognition of these emotions by listeners (see Juslin and Laukka, 2003 and Scherer and Bänziger, 2004 for reviews). However, vocal changes are not always symptoms of the speaker's internal emotional state; for example, prosody can be intentionally used as an expressive device to elicit empathy in the listener, allowing interactants to experience (Aziz-Zadeh et al., 2010; Rodero, 2011) and understand (Regenbogen et al., 2012; Ong et al., 2018) the speaker's feelings. This combination of affective and inferential processing of prosody provides the speaker with important emotional influence and bolsters supportive behavior from the listener, with potential social benefits for both parties (Van Kleef, 2009; Pell and Kotz, 2020).

It has been reported that prosodic features of complaints signal increased affectivity through elevated mean fundamental frequency and frequency variability, syllable elongation, and emphatic accentuations (Acuña-Ferreira, 2002; Ogden, 2010; Selting, 2010; Rao, 2013). In emotional contexts, these acoustic changes are often associated with negative and high arousal emotions, like anger, sadness, surprise and indignation (Drew, 1998; Selting, 2010). Complaints may also be viewed as expressions of pain and suffering, which are associated with specific forms of vocal expression (Lerner et al., 2016; Raine et al., 2019). The present study is based on a large set of complaining utterances that display many of the acoustic and emotional properties described above, as well as voice quality patterns that resemble expressions of simulated pain (Mauchand and Pell, n.d. a, under review; Raine et al., 2019).

While the role of prosody in communicating the emotive involvement of complainers is heavily suggested, most of the literature on complaints comes from the pragmatics field, based

largely on descriptive and qualitative analyses of conversations (Boxer, 1993; Drew, 1998; Acuña-Ferreira, 2002; Edwards, 2005; Traverso, 2009; Ogden, 2010; Selting, 2010; Rao, 2013). No experimental investigation has been conducted to establish how prosody affects the perception of complaints, especially with respect to other lexical or contextual cues that complainers often provide. As mentioned above, the emotive involvement of the speaker is often more important than the object of the complaint, meaning that even innocuous topics can be the focus of valid complaints (Boxer, 1993). Still, the preference of complainers to provide specific descriptions (Alicke et al., 1992), expletives (Drew, 1998), and extreme-case formulation (Selting, 2010) suggest that complaining emotive interactions depend on both linguistic and paralinguistic cues, albeit in an unclear manner.

The integration of prosodic and verbal affective signals and their combined impacts on social perception can be complex. The relative effects of cues in each channel may vary at different stages of perception, processing, and evaluation (Paulmann and Kotz, 2008; Pell et al., 2011; Meconi et al., 2018), and likely depend heavily on task demands (Regenbogen et al., 2012) and the emotional salience of cues (Wambacq and Jerger, 2004). In expressive speech acts, the role of prosody is traditionally described as an indirect, illocutionary force that can only convey meaning with the appropriate verbal statement (Grice, 1989; Wichmann, 2000). However, recent studies suggest that prosody alone can reveal the intentions of a speaker in a powerful manner (Hellbernd and Sammler, 2016; Caballero et al., 2018; Truesdale and Pell, 2018). For example, in motivating and persuasive speech, prosody can "tag" verbal information as important and increase the persuasiveness of a speaker even when the verbal information is not credible (Zougkou et al., 2017; Van Zant and Berger, 2020). Prosody is thus an important emotive *and* persuasive device in low-involvement communicative situations (Gelinas-Chebat and Chebat, 1992), which is often the case of third-party complaints (Alicke et al., 1992; Boxer, 1993).

The use of affect as social information further depends on a number of social-relational factors, such as cultural display rules, familiarity, or group biases (Van Kleef, 2009). Indeed, if the traditional view of emotions as genuine biological responses could imply a universal consistency in their expression (Frick, 1985; Ekman et al., 1987), describing affective displays as social tools implies investigating how social and cultural contexts affect their usage (Van Kleef, 2009; Scarantino, 2017). Several studies already suggest that despite a basic universality, emotional communication can be affected by cultural in-group advantages (Elfenbein and Ambady, 2002), depend on cultural proximity (Laukka et al., 2016) and seem to mainly affect positive rather than negative emotions (Sauter et al., 2010; Scherer et al., 2011; see Laukka and Elfenbein, 2020 for a review). Often, out-group accent perception does not impede how well emotions are recognized but does affect perceived intensity, empathic arousal or physiological responses from listeners (Soto and Levenson, 2009; Mac et al., 2010; Thierry et al., 2015). Beyond emotions, a speaker's accent is a marker of identity: the information (or lack thereof) that it carries is known to interfere with speech processing (Floccia et al., 2006; Sumner and Samuel, 2009), create biases and stereotypes (Kuiper, 2005; Lev-Ari and Keysar, 2010; Heblich et al., 2015), and affect the appraisal of diverse pragmatic cues (Yuan et al., 2019; Jiang et al., 2020). Cultural factors may thus affect numerous stages of production, perception, and interpretation of emotive speech.

Complaining appears to be a convention rooted in a number of cultures. Be it the French *se plaindre* (Traverso, 2009), the Australian *whinge* (Edwards, 2005), the German *Jammern* (Winchatz, 2016), or the Israeli *kiturim* (Katriel, 2013), many societies have defined complaining as a cultural custom, each with their own specificities and social implications. These potential cultural specificities raise the question of what constitutes a complaint across cultures. Yet, few studies have directly investigated the cross-cultural aspect of complaints. An investigation by Rao (2013) reported that Mexican Spanish complaints showed intonational variation typical of European Spanish complaints, but in a more accentuated manner. Similarly, Mauchand and Pell, n.d. b, under review reported that Canadian French (Québécois) and European French complaints show strong acoustic similarities but sometimes differ in the weight given to certain prosodic cues and the emotional representations they convey. Parallel work on direct complaints also show some pragmatic differences between native and non-native complaints (Trenchs, 1994; Kraft and Geluykens, 2002). Beyond the definition of complaints, these acoustic differences could affect cross-cultural understanding of complaining speech, individuals being potentially more sensitive to emotive prosodic signals from their own group. To date, work which sheds light on these questions has not been undertaken.

The goal of the present study was to give insight on how third-party complaints are perceived from affective prosody and other cues that mark a speaker's "complaining intentions," using the Emotions as Social Information (EASI) model as a general reference (Van Kleef, 2009). Furthermore, we explored the role of social-relational variables in this context by studying two francophone groups: French (i.e., European French) and Québécois (i.e., French Canadian). While mutually intelligible, these two groups have different cultural backgrounds and distinct accents, thus allowing the isolation of cultural group (dis)advantages in the processing of complaints in the absence of language barriers. French and Québécois participants listened to pre-validated utterances that varied in prosody, verbal content, and speaker accent, and evaluated "how complaining" each utterance sounded. The study also investigated the relationship between encoding and decoding processes by analyzing how the perception of expressive speech acts, such as complaints, is driven by particular acoustic features of vocal affect signals. It was predicted that a speaker's tone of voice would be the main marker of a complaining intention, especially when verbal cues did not convey high emotive involvement, i.e., when speakers complained about innocuous rather than explicitly pain-related topics. The detection of complaints was expected to depend on how the speakers produced emotive signals, especially through modulation of voice pitch and other emotion-related cues, which are likely to mediate the effect of complaining prosody on participant's evaluations. Finally, it was predicted that social-relational factors would influence complaint perception: participants were expected to discriminate complaints from

neutral utterances better for speakers of their own cultural group, potentially because of underlying biases and/or specific display rules associated with complaining speech.

MATERIALS AND METHODS

Participants

Power analyses for mixed models (Judd et al., 2016) were performed to determine the required number of participants. Large effects of prosody and verbal content were reported in previous studies with similar procedures (Caballero et al., 2018; Mauchand et al., 2020). Due to the large number of stimuli ($n = 320$), less than 25 participants were required to attain power over 99% for these effects. The effect of culture, if present, would be smaller based on previous cross-cultural studies that have used recognition tasks (Elfenbein and Ambady, 2002; Liu et al., 2015; Laukka et al., 2016; Jiang et al., 2018). Based on an effect size of 0.3 with intercepts and slopes variances of 0.1, a minimum sample size of 57 participants would be required to achieve power of 90% for this variable.

In total, 31 French and 27 Québécois participants, aged 18–35, with no hearing or neurological impairment were recruited in the Montréal area. French participants were born in France, had lived in France until at least 18, and had arrived less than 3 years ago in Montréal (for study or work). Québécois participants were born and lived in Québec (a French-speaking province in Canada) until age 18 and had never lived in France or another francophone country. All participants spoke French as their mother tongue.

Data about participants' personality and cultural attitudes were collected through a number of tests and questionnaires (see Mauchand and Pell, n.d. a, under review for a full report on these measures). Accent-based implicit biases were measured through a modified Implicit Association Test (Greenwald et al., 1998) consisting of Pleasant and Unpleasant words presented together with French and Québécois neutral utterances (Mauchand and Pell, n.d. a, under review). Explicit attitudes toward French and Québécois populations were probed through a questionnaire based on the Stereotype Content Model (Fiske et al., 2002), composed of 20 questions about the perceived Warmth and Competence of each community. Finally, empathic abilities were assessed through the Perspective-Taking and Empathic Concern subscales of the French version of the Interpersonal Reactivity Index (Gilet et al., 2013).

Materials

Materials were created and validated in a previous study focusing on the acoustic dimensions of speech complaints (Mauchand and Pell, n.d. b, under review). Stimuli were short spoken utterances describing a past event, constructed in the form of *token sets* (each composed of 4 unique utterances). A token set was built around a root sentence that was manipulated in two ways. First, we modified the verbal content by modifying the last word of the statement, to refer to a neutral event, e.g., "Il a dit que j'étais sorti/He said I was outside" (*Control* condition) or a socially painful event for the speaker, e.g., "Il a dit que j'étais stupide/He said I was stupid" (*Pain* condition). The list of sentences, together with their English translation, can be found in the **Appendix**. For each type of statement, we then manipulated the form of prosodic expression: speakers uttered each sentence in a manner as if simply reporting the event (*Neutral* condition) or as if complaining to a friend (*Complaint* condition). One token set was thus composed of 4 utterances with different Statement/Prosody combinations: *Control/Neutral, Control/Complaint, Pain/Neutral, Pain/Complaint.* Forty-two token sets were thereby created.

Initially, 672 utterances were produced by 4 French and 4 Québécois speakers (2 males and 2 females in each group) in order to modulate accent/sociocultural features of the stimuli. Recordings were digitally captured in a sound-attenuated chamber with a high-quality head-mounted microphone onto a Tascam recorder (sampling rate of 44.1 kHz, 16-bit, mono, wav format). They were then edited in Praat (Boersma and van Heuven, 2001) into short.wav audio files and normalized to a peak intensity of 70 dB.

A short validation study was conducted to ensure the quality of the recordings and to select a subset of the stimuli for the current study. Ten French (5 males, 5 females, age: $M = 21.1$, sd = 3.8) and 9 Québécois (3 males, 6 females, age: $M = 23.00$, sd = 2.78) participants listened to all utterances from their own group ($n = 336$) and evaluated: (1) whether an utterance was a complaint (yes/no); and (2) if it was a complaint, its intensity of expression on a 5-point scale. Results of the validation task are displayed in **Table 1**. Pain/Complaint utterances were almost unanimously considered complaints with high intensity ratings, while Control/Neutral utterances were very rarely considered complaints. Results for Pain/Neutral and Control/Pain utterances suggest that prosody had a larger impact than statement type on the perception of complaints.

For the present study, a subset of utterances was selected to minimize the repetition of sentences in the experiment, to remove

TABLE 1 | Results of the validation/selection task, by speaker group (mean + standard deviation).

Utterance/ prosody type	Unselected utterances ($n = 352$)		Selected utterances ($n = 320$)	
	Proportion of YES answers to "is the person complaining?"	Intensity rating[a] (1 to 5)	Proportion of YES answers to "is the person complaining?"	Intensity rating[a] (1 to 5)
Québécois				
Control/Neutral	0.26 (0.27)	1.80 (0.87)	0.21 (0.20)	1.89 (0.99)
Control/Complaint	0.81 (0.21)	2.94 (0.91)	0.81 (0.19)	2.89 (0.70)
Pain/Neutral	0.60 (0.21)	1.74 (0.63)	0.64 (0.21)	1.69 (0.63)
Pain/Complaint	0.96 (0.07)	3.15 (0.84)	0.98 (0.05)	3.34 (0.57)
French				
Control/Neutral	0.23 (0.16)	1.53 (0.54)	0.25 (0.15)	1.43 (0.45)
Control/Complaint	0.86 (0.13)	3.45 (0.78)	0.87 (0.13)	3.37 (0.62)
Pain/Neutral	0.57 (0.16)	1.64 (0.38)	0.53 (0.18)	1.67 (0.43)
Pain/Complaint	0.95 (0.06)	3.47 (0.60)	0.94 (0.08)	3.46 (0.52)

[a]*Note that the rating is only made when answering YES to the previous question.*

potential outliers, and to ensure that stimuli were representative of the speakers' intentions (complaining vs. neutral) according to listeners from their own group. For each speaker, a token set was selected if there was enough consensus that the Control/Neutral utterance was NOT a complaint and that the Pain/Complaint utterance was indeed a complaint with high intensity ratings. To avoid selection bias on the prosody/statement effects, results for "incongruent" utterances were not taken into account for the selection. Moreover, each speaker had a "mirror" speaker (of the same sex) in the other cultural group that uttered exactly the same token sets, such that each token set was present exactly once in each group. This selection process did not affect the overall perceptual quality of the stimulus set, as scores for selected and unselected items remained close. In total, there were 2 Accents × 4 Speakers × 10 Token Sets × 2 Prosodies × 2 Statements = 320 selected utterances.

Acoustic measures for each of the 320 selected utterances were collected using the Geneva Minimalistic Acoustic Parameter Set/GeMAPS (Eyben et al., 2016) package from the publicly available openSMILE toolkit (Eyben et al., 2010). The GeMAPS constitutes a reliable standardized baseline set of affect-related acoustic measures (for more details on the computation and implementation of the measures, see Eyben et al., 2016). A full acoustic analysis of all 672 stimuli is presented in a previous study (Mauchand and Pell, n.d. a, under review). The present study focuses on measures of pitch, voice quality, and rhythm known to be perceptually relevant in complaint production (Acuña-Ferreira, 2002; Rao, 2013) and other related modes of emotional expression (Laukka et al., 2016; Raine et al., 2019). Note that since the volume of stimuli was normalized for perception, intensity-related acoustic measures could not be reliably extracted for consideration in the present study. The following acoustic measures were computed as a mean measure over the full duration of each utterance:

- F0, the fundamental frequency, indexing pitch on a logarithmic semitone scale. Considering the importance of pitch in complaints, both the mean (F0 M) and the rescaled standard deviation (F0SD) over the utterance were computed.
- Jitter, indexing aperiodicity (instability) of the F0 signal – voice "creakiness"
- Shimmer, the difference of the peak amplitudes of consecutive F0 periods, indexing voice roughness in dB
- Harmonics-to-Noise Ratio (HNR), indexing the relative amount of additive noise in the voice
- F1, first formant center frequency in Hertz
- Utterance duration and final word duration in seconds (computed on Praat).

These measures are summarized in **Table 2**.

TABLE 2 | Summary of acoustic measures for the selected stimuli for each speaker group (Mean + standard deviation).

	French		Québécois	
	Neutral	**Complaint**	**Neutral**	**Complaint**
F0 M	28.43 (4.43)	34.13 (3.57)	28.93 (5.23)	34.28 (6.14)
F0 SD	0.14 (0.06)	0.15 (0.08)	0.17 (0.09)	0.20 (0.08)
HNR	6.64 (2.4)	9.22 (1.66)	6.40 (2.59)	7.83 (2.94)
Jitter	0.05 (0.02)	0.04 (0.02)	0.06 (0.03)	0.05 (0.03)
Shimmer	1.31 (0.42)	1.06 (0.35)	1.35 (0.44)	1.18 (0.34)
F1	535.98 (87.65)	554.86 (74.77)	534.44 (61.28)	559.84 (60.59)
Duration	1.16 (0.26)	1.30 (0.30)	1.42 (0.29)	1.46 (0.39)
Final word duration	0.38 (0.12)	0.45 (0.14)	0.47 (0.15)	0.56 (0.17)

F0 M, mean fundamental frequency; F0 SD, rescaled standard deviation of the fundamental frequency; HNR, Harmonics-to-Noise Ratio; F1, first formant center frequency.

Procedure

Each participant was presented all 320 selected stimuli in a fully randomized order using Cedrus Superlab 5 software. The stimuli were divided in 8 blocks of 40 utterances, with a self-monitored break between each block. After presentation of an utterance, participants answered the question "À quel point cette personne est-elle en train de se plaindre?" (*How much is this person complaining?*) on a 7-point Likert scale ranging from *Pas du tout* (*Not at all*) to *Énormément* (*Very much*) by pressing a button on a response box. No time limit was set. Participants were not given any indication or strategy on how to form their answer and were told that there was no right or wrong answer. The whole experiment lasted a little more than 1 h.

RESULTS

Main Model

Participant's ratings were analyzed through a Linear Mixed Effect Model using the R package lme4 (Bates et al., 2015). *T*-tests and *p*-values were computed with Satterthwaite's approximation using lmerTest package (Kuznetsova et al., 2017). The model was built with the participant's Response (0–6) as the response variable, and Participant Culture (French/Québécois), Speaker Accent (French/Québécois), Statement (Control/Pain), and Prosody (Neutral/Complaint) as predictors. All 2- and 3-way interactions were also entered as predictor terms. Participant and Speaker/TokenSet were added as random intercepts: TokenSet was nested within Speaker, such that Speaker was one random intercept and the interaction between Speaker and TokenSet was another random intercept, thus accounting for the variability of speakers and the variability of token sets within each speaker. Additionally, Culture, Statement and Prosody were added as uncorrelated by-Speaker/TokenSet slopes, and Accent, Statement and Prosody were added as uncorrelated by-Participant slopes.

The model accounted for 70% of the variance in the data ($r^2 = 0.70$). The model revealed a significant effect of Content ($\beta = 0.91$, se $= 0.13$, $t(19.13) = 7.04$, $p < 0.001$), suggesting that when speakers provided linguistic evidence of a painful situation (Pain vs. Control statement), ratings increased by almost 1 point on the scale. A larger effect of Prosody was observed ($\beta = 2.38$, se $= 0.21$, $t(12.82) = 11.38$, $p < 0.001$); statements expressed

in a complaining versus neutral tone tended to increase ratings by more than 2 points. Speaker accent was associated with a marginal, yet noticeable effect, as statements produced in the Québécois accent tended to be rated stronger exemplars of complaints than those produced in the French accent ($\beta = 0.68$, $se = 0.30$, $t(6.20) = 2.27$, $p = 0.063$). This trend was informed by another marginally significant pattern in the data, representing a 3-way interaction of Participant Culture, Speaker Accent and Prosody ($\beta = 0.35$, $se = 0.18$, $t(13.71) = 1.98$, $p = 0.068$). The effects of Prosody (Complaint > Neutral) on complaint perception tended to be greater when Québécois participants were listening to the Québécois accent. No other term showed a significant effect (ps > 0.1). Results are summarized in **Table 3**; Content and Prosody effects are detailed in **Figure 1**.

Follow-up analyses were run to further investigate the relative effects of lexical and prosodic manipulations on the participant's responses. Looking at the model's random slopes reveals important variance in these two predictors (0.82 for the slope of Prosody by Participant, 0.45 for the slope of Statement by Participant). A large negative correlation between the two slopes was found ($r = -0.52$), indicating that participants with greater Prosody coefficients tended to have smaller Statement coefficients (see **Figure 2**). Note that a possible outlier showing extreme coefficients can be seen on **Figure 2** but removing this participant from the analysis did not affect results. Correlations were then calculated between the predicted random effects of Prosody and Content and IRI scores, revealing a medium correlation between a participant's predicted Prosody effect and their score on the Perspective Taking scale ($r = 0.21$), but not the Empathic Concern scale ($r = 0.06$). This pattern was mirrored in correlations with the predicted Statement effect, although to a much lesser extent (PT scale: $r = -0.13$; EC scale: $r = -0.01$). These results suggest that participants who were more sensitive to complaining prosody (especially those with greater perspective-taking skills) relied less on the actual statements.

Analyses also probed the effect of speaker accent and assessed whether this effect could be driven by more specific social-relational factors such as cultural attitudes. A measure of implicit cultural bias was derived from a customized version of the IAT, designed to measure implicit attitudes held by French and Québécois toward speakers of each group (Greenwald et al., 1998; Mauchand and Pell, n.d. a, under review) Based on the IAT D-score, the predicted random effects for each participant showed no particular relationship with implicit biases toward speakers of each cultural group ($r = 0.04$) nor with any of the Stereotype Content Model scores for either group (French Warmth: $r = -0.05$; French Competence: $r = 0.04$; Québécois Warmth: $r = 0.10$; Québécois Competence: $r = 0.07$). This suggests that accent effects were not strongly driven by implicit or explicit cultural biases.

TABLE 3 | Mean rating of "how much the speaker is complaining" by French and Québécois listeners, according to the speaker's accent, prosody, and the type of statement (0–6 scale).

			French participants	Québécois participants
Accent	**Statement**	**Prosody**	**M (SD)**	**M (SD)**
French	Control	Neutral	0.92 (0.66)	0.89 (0.63)
		Complaint	3.28 (0.85)	3.27 (1.03)
	Pain	Neutral	1.94 (0.95)	1.70 (1.10)
		Complaint	3.96 (0.74)	3.91 (0.79)
Québécois	Control	Neutral	1.32 (0.72)	1.22 (0.77)
		Complaint	3.88 (0.85)	4.22 (1.07)
	Pain	Neutral	2.72 (0.85)	2.37 (1.01)
		Complaint	4.76 (0.62)	4.90 (0.62)

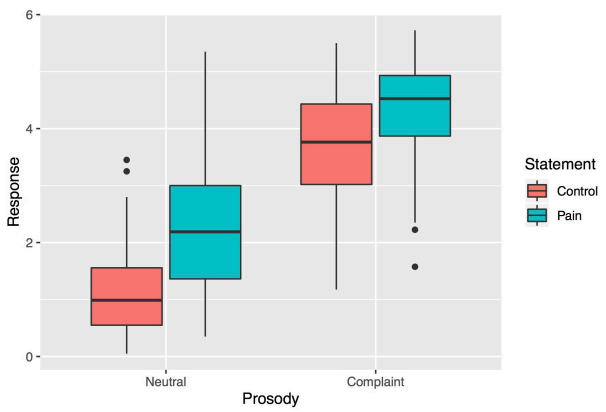


FIGURE 1 | Box-plot summary of Prosody and Statement effects, averaged by participant.

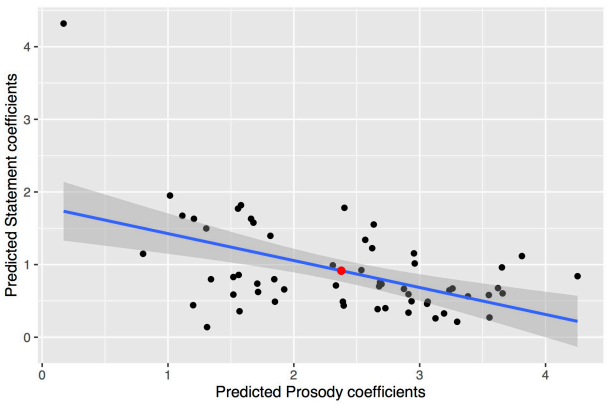


FIGURE 2 | Relationship and linear regression slope between each participant's predicted coefficients for Statement and Prosody effects in the linear mixed-effects model. The red point indexes the actual coefficients from the whole model.

Mediation of Prosody Effects by Acoustic Parameters

To determine how the effect of prosody on complaint perception relates to specific acoustic properties of the voice, a regression-based mediation analysis with multiple mediators was run following Vanderweele and Vansteelandt (2014). This method

accounts for potential relationships between mediators and prevents any effect overlap and redundancies of running several individual mediation analyses. Acoustic parameters described in the Methods section were selected as mediators (see **Table 2**). The measures from each utterance were standardized by subtracting the mean and dividing by the standard deviation of all utterances. The mediation analysis was thus performed with Prosody as the treatment variable, Response as the outcome variable, and the eight acoustic parameters as mediators.

First, to assess how the treatment variable Prosody affected the mediators, eight linear regressions were run, each with a mediator as the response variable and Prosody as the predictor. Then, to evaluate the effects of the treatment and mediators on the outcome, a multiple linear regression was run with Response as the response variable and the treatment (Prosody) and all eight mediators as predictors. The direct effect of Prosody is given by its coefficient in the latter regression model; the indirect effect of Prosody through a given mediator is given by the product of the effect of Prosody on this mediator and the effect of the mediator on the Response; the total indirect effect of prosody is given by the sum of all such mediated effects.

The speaker's mode of prosodic expression had significant effects on each mediator: compared to neutral statements, complaints showed increased F0M ($\beta = 0.49$, se < 0.01, $t = 76.79$, $p < 0.001$), increased F0SD ($\beta = 0.11$, se < 0.01, $t = 15.39$, $p < 0.001$), reduced shimmer ($\beta = -0.26$, se < 0.01, $t = -37.16$, $p < 0.001$), reduced jitter ($\beta = -0.06$, se < 0.01, $t = -8.39$, $p < 0.001$), longer utterance duration (($\beta = 0.09$, se < 0.01, $t = 17.91$, $p < 0.001$), and final word duration ($\beta = 0.25$, se < 0.01, $t = 34.56$, $p < 0.001$), increased HNR ($\beta = 0.38$, se < 0.01, $t = 55.22$, $p < 0.001$), and increased F1 ($\beta = 0.15$, se < 0.01, $t = 21.09$, $p < 0.001$). In turn, participant's Response/ratings were positively affected by F0M ($\beta = 1.62$, se $= 0.05$, $t = 29.52$, $p < 0.001$), Jitter ($\beta = 0.25$, se $= 0.03$, $t = 7.49$, $p < 0.001$), and utterance duration ($\beta = 0.20$, se $= 0.03$, $t = 6.51$, $p < 0.001$), and negatively affected by Shimmer ($\beta = -0.16$, se $= 0.04$, $t = -4.22$, $p < 0.001$), HNR ($\beta = -0.78$, se $= 0.06$, $t = -12.77$, $p < 0.001$), and F1 ($\beta = -0.19$, se $= 0.03$, $t = -6.20$, $p < 0.001$. No effect of F0SD ($\beta < 0.06$, se $= 0.03$, $t = 1.67$, $p = 0.094$) or final word duration ($\beta < -0.02$, se $= 0.03$, $t = -0.81$, $p = 0.420$) were found. As shown in **Figure 3**, F0M was by far the greatest mediator of Prosody on Response ($\beta = 0.79$), followed by shimmer ($\beta = 0.04$) and utterance duration ($\beta = 0.02$). Meanwhile, the mediations of HNR ($\beta = -0.29$), F1 ($\beta = -0.03$), and jitter ($\beta = -0.02$) were negative. Most of the Prosody effect was not linearly mediated by acoustic measures, as the total indirect effect of Prosody ($\beta = 0.54$) accounted for much less of the total effect ($\beta = 2.38$). The mediation model is illustrated in **Figure 3**.

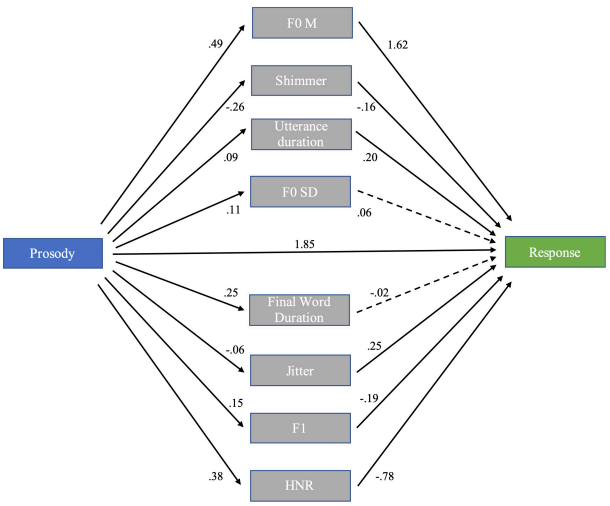


FIGURE 3 | Model summary for the multiple mediation analysis. F0M, mean fundamental frequency; HNR, Harmonics-to-Noise Ratio; F0SD, standard deviation of the fundamental frequency.

DISCUSSION

Our results provide experimental evidence supporting the literature on complaints, emotive communication, and vocal affect. As elaborated below, they emphasize the important role of prosody in conveying emotive information in communication and its relationship to other message-level (e.g., lexical) and social-relational (e.g., cultural) dimensions of social interaction.

Affective Prosody, Effective Complaint

The core of the study measured how listeners evaluate the complaining nature of utterances in different situations combining several factors. The manipulation of the speaker's prosody was revealed to have the largest effect on listener's evaluations; everything else controlled for, switching from a "neutral" to a complaining tone of voice led to a marked increase in whether statements were judged to be a complaint. This finding parallels the ability to recognize basic emotions and evaluate speaker arousal from vocal expressions (Scherer, 2003; Grandjean et al., 2006), extending this ability to the general perception of a speaker's emotivity in discourse. Through the speaker's intention to foreground speaker affect (Arndt and Janney, 1991; Caffi and Janney, 1994), our results show that complaints can be discriminated from vocal signals without requiring complex attitudinal inferences from situational context (Wichmann, 2000). This exemplifies the use of emotional expressions as a social tool, providing listeners with affective information that allows them to make inferences and to (voluntarily) share the speaker's emotive state (Van Kleef, 2009; Scarantino, 2017). Here, prosody appears to be the main device in the collaborative treatment of affectivity (Drew, 1998; Selting, 2010), constituting a relatively direct and effective way for listeners to assess a complainer's subjective state (Acuña-Ferreira, 2002; Edwards, 2005). It can be said that detecting the emotivity of the speaker is a crucial first step in complaining interactions; by allowing listeners to recognize complaints, prosody is likely to play a key role in facilitating the affiliative and empathic response of listeners (Boxer, 1993; Traverso, 2009).

Prosody was not the only way speakers could influence listener's evaluation of complaints. Utterances that described an

explicitly painful situation were perceived as stronger exemplars of complaints than statements which did not. Affective words and sentences are known to affect a listener's perception of emotions in speech (Pell et al., 2011; Regenbogen et al., 2012; Rigoulot et al., 2020). However, this effect did not interact with prosody and was small enough that control statements spoken in a complaining tone were perceived as more complaining than pain-related sentences in a neutral tone. This confirms an important characteristic of complaints: *how* we complain is more important than *what* we complain about, and one can virtually complain about anything (Alicke et al., 1992; Boxer, 1993). Still, the description of a past situation that would typically be associated with (social) pain can facilitate the perception of an utterance as a complaint; this factor is likely to play a role in how complaining interactions unfold in spontaneous interactions.

Interestingly, the perceptual weight given to the statement seemed to be greater when prosody was less efficient; listeners who were less sensitive to prosodic signals could presumably compensate by relying on the more tangible, explicit nature of verbal information (Zougkou et al., 2017). The relative weighting of prosodic and linguistic information can be partially explained by listener's empathic abilities; individuals with heightened perspective-taking skills (or *cognitive* empathy), relied more on prosody and less on the verbal statement. In contrast, participants with greater empathic concern (or *affective* empathy) did not show such associations. These results are congruent with the nature of the task, which required *understanding* the speaker's intention; in this context, the interpretation of the displayed affect would have been driven by inferential rather than affective processes (Van Kleef, 2009). Future research using other designs such as self-ratings or physiological measures (de Vignemont and Singer, 2006; Lang et al., 2011; Kanske et al., 2015) could further distinguish affective from inferential processes in empathy and assess how listeners actually *share* a complainer's affective state from prosody.

From Acoustic Signals to Emotive Representations

The manipulation of prosody in our study allowed us to determine to what extent these cues are instrumental for listeners to recognize the speaker's intent to complain; however, it does not explain *which* acoustic cues drive these judgments and *how* they do it. Prosody researchers who have adopted a Brunswikian approach have stressed that while emotion encoding and decoding have been widely covered by the literature in a separate manner, investigations that *combine* both processes are lacking (Juslin and Laukka, 2003; Scherer, 2003; Grandjean et al., 2006). Acoustic analyses of the present stimuli had revealed a number of parameters that speakers seem to manipulate in order to convey their complaints (Mauchand and Pell, n.d. b, under review). In particular, increased mean F0 and F0 variability, decreased shimmer, increased Harmonics-to-Noise ratio, and lengthened final word were widely used acoustic strategies to communicate complaints. The multiple mediation analysis performed here assessed if and how these parameters were actually used by listeners in their evaluations.

Results of the mediation analysis suggest that mean F0 was by far the most important acoustic parameter in mediating the effect of Prosody; complainers increased their mean pitch, which was perceived as more complaining by listeners. Fundamental frequency is known to be the most directly accessible marker of affect for listeners, and is modulated in both a discrete and continuous manner to express basic emotions (Frick, 1985; Grandjean et al., 2006; Eyben et al., 2016) and attitudes (Jiang and Pell, 2017; Caballero et al., 2018; Mauchand et al., 2018; Truesdale and Pell, 2018). Increased F0 mean also marks non-aggressivity and is central to affiliative behaviors as described by the Frequency code (Ohala, 1984; Gussenhoven, 2004), which could explain its central importance in the production and perception of complaints.

Differences in voice quality also showed notable patterns in mediating the effect of prosody on complaint recognition. Compared to neutral speech, complaints displayed reduced shimmer, increased HNR, and to a lesser extent reduced jitter, indicating that speakers employed a less rough, less creaky and less noisy voice when they were complaining. Evidence of increased voice control (Latoszek et al., 2018) while complaining is also characteristic of simulated but not natural pain (Lautenbacher et al., 2017; Raine et al., 2019). Interestingly, HNR and Jitter negatively mediated the participant's response, suggesting that listeners may perceive that complaints are not genuine but reconstructed displays of affect (Selting, 2010). This impression may also explain why even complaints with pain-related statements were rarely evaluated using the highest points on the scale. In addition, reduced shimmer was associated with a slight increase in complaint ratings, possibly due to the importance of this acoustic marker in detecting sadness (Juslin and Laukka, 2003). Increased F0SD and Final World Duration, which were associated with complaining prosody, did not significantly affect listener's judgments in the current study. It should be borne in mind that complaints occur in complex interactions, and the role of prosodic features may not be limited to signaling an emotive intent. Dynamic variations in pitch and rhythm, which mark the emphatic structure of speech (among others), could instead help to coordinate the upcoming interaction and indicate how the collaborative treatment of affect should proceed (Selting, 1994; Szczepek Reed, 2011). Also, the fact that effort-related parameters, such as higher F1 and larger F0 variation (Traunmüller and Eriksson, 2000), had little or even negative effects on the perception of complaints reaffirms that successful complaints are conveyed through affiliative signals (as per the Frequency code), rather than effort-derived meanings (as per the Effort code) (Ohala, 1984; Gussenhoven, 2004).

It is important to note that while a portion of the prosody effect on complaint perception was mediated by specific forms of acoustic change, a large part of the effect remains unexplained in the model. As our selected acoustic parameters cover many of the core acoustic features of utterances (except loudness), it is unlikely that entering more parameters as mediators would significantly increase the proportion of the mediated effect. Instead, it appears that the transformation of acoustic signals into an emotive representation is not

a linear process that can be fully decomposed. In the context of our task, it is likely that the apparent contrast between neutral and complaining prosody allowed a discrete categorization of the two utterance types; the relative salience of certain parameters (such as pitch or vocal noise) could then further modulate the perception of utterances within each category.

Social-Relational Factors in Emotive Communication

While evaluations of complaints relied mainly on prosodic and lexical information, the cultural manipulation of this experiment had a marginal, but still noteworthy, impact on perceptual judgments. Overall, Québécois speakers were rated as producing stronger (i.e., more prototypical) complaints than French speakers, and there was a strong trend for Québécois listeners to recognize complaining prosody better when produced by other Québécois speakers.

In a previous study (Mauchand and Pell, n.d. b, under review), differences between French and Québécois complaints were reported at both the acoustic and perceptual level, motivating our continued interest in how socio-cultural variables influence complaint perception. In that study, we found that Québécois speakers, when complaining, used greater pitch variability and distinct rhythmic patterns than French speakers and were perceived as angrier and more surprised (as opposed to sad for the French speakers, Mauchand and Pell, n.d. b, under review). Of key interest, Québécois speakers used a harsher voice quality than French speakers when producing complaints (reduced HNR). Here, the mediation analysis revealed that HNR reduced the intensity of the perceived complaints; the harsher vocal quality of Québécois speakers might thus have facilitated the detection of complaints by certain listeners. This facilitation was enhanced at the in-group level, as Québécois listeners seemed more attuned to prosodic contrasts produced by other Québécois speakers. This finding suggests the existence of cultural display rules and in-group advantages in emotive speech communication as is the case for the expression of emotions (Elfenbein and Ambady, 2002). However, the absence of a similar in-group advantage for the French group suggests this effect might depend on the interplay of individual, cultural and contextual factors. For example, the exposure of our French participants to the Québécois culture in this study could have reduced potential in-group advantages for that group. However, French participants were very recent immigrants in Québec, and most of them reported having very few Québécois people in their social and professional circles. On the other hand, Québécois participants reported having more French acquaintances, and are frequently exposed to French-accented speech from an early age (Kircher, 2012). Thus, the asymmetry in cultural effects may alternatively be due to a lack of sensitivity of French participants to the more expressive Québécois complaining style.

While the decoding of emotive cues in the voice may be enhanced for certain in-group interactions, this facility does not seem to originate from cultural bias or prejudice. No relationship was found between the effect of accent and either implicit or explicit biases toward either group, even though such biases exist between French and Québécois communities (Auger and Valdman, 1999; Kircher, 2012; Mauchand and Pell, n.d. a, under review). While stereotypes and prejudice do affect neural activity (Quadflieg and Macrae, 2011; Jiang et al., 2018) and affective empathy (Xu et al., 2009; Contreras-Huerta et al., 2014), they often don't impede speech comprehension and affect recognition (Gill, 1994; Lev-Ari and Keysar, 2010; Thierry et al., 2015). Thus, accent effects may instead arise from processing issues and/or differential use of prosodic signals. Even then, the potential impact of accent cues were minimal when compared to the efficacy of both speaker groups to convey a complaining intention through prosody. These results thus reveal a strong consistency of speakers in intentionally using emotions as social signals and of listeners to infer their intentions in the case of complaints. This inference process can be subtly modulated by social-relational factors, such as the culturally normative usage of certain prosodic cues (Elfenbein and Ambady, 2002; Laukka et al., 2016; Scherer, 2003; Van Kleef, 2009). Other factors not taken into account here may also play an important role in natural complaint perception: here, the absence of context, visual cues, or a true indication of the social proximity between speaker and listener might explain why evaluations of complaints rarely reached the end of the scale. Sex/gender is also often mentioned as an important factor in complaining (Acuña-Ferreira, 2002; Selting, 2010); anecdotally, speaker sex was tentatively added as a parameter in our model, but did not show any significant effect (although this could be due to the small number of male/female speakers in our experiment). Future studies should investigate how a wider range of these social factors influence inferential and affective processes underlying emotive speech communication.

CONCLUSION

As the first perceptual investigation of complaining speech, the present study reaffirms the central role of prosody as a social device to foreground the emotive state of the speaker. The effective production and appraisal of emotive features in the voice denote a tacit understanding between speaker and listener on how complaints are performed, which depends on the capacity of listeners to detect these signals and collaborate with the social goals of the speaker (i.e., to commiserate and co-complain). Listeners also use linguistic evidence describing the nature and/or antecedents of a complaint when evaluating these speech acts, although these cues may be less diagnostic than prosodic contrasts for determining when a speaker intends to complain (and seek social affiliation and support). As such, complaints can be qualified as acts of manipulation without deception, similar to other emotive acts like persuasion, motivation or charismatic speech: intentional displays of emotion that regulate the dispositional affect of listeners and promote social affiliation. This metapragmatic understanding of human affect, central to speaker/listener relationships, needs to be systematically considered in future investigations of speech, attitudes, and emotions (Pell and Kotz, 2020). Including social-relational factors, such as cultural relationships, is crucial to advance

perspectives in this literature; future work should investigate how more distant cultures communicate complaints and other types of emotive meanings. Experimental approaches that include empathic assessments, neurophysiological measures, or which study group interactions would also produce valuable evidence to build on theoretical frameworks describing emotive communication, affect, and prosody.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. Data and materials can be found on the OpenScience Framework (Foster and Deardorff, 2017; https://osf.io/9az68/?view_only=9a5529fc1a0645aa9b17f6589a79e848).

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Faculty of Medicine and Health Sciences Institutional Review Board (McGill IRB). The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

MM designed the experiment, collected, analyzed and interpreted data, and drafted and developed the manuscript. MP supervised the experiment and provided critical revisions of the manuscript. Both authors contributed to the article and approved the submitted version.

FUNDING

This research was supported by an Insight Grant to MP from the Social Sciences and Humanities Research Council of Canada (G245592 SSHRC 435-2017-0885) as well as scholarships to MM from McGill Faculty of Medicine, the Integrated Program in Neuroscience at McGill, and the Quebec Bio-Imaging Network (QBIN/RBIQ).

ACKNOWLEDGMENTS

The authors would like to thank Emmanuelle Beauvais-Lacasse for her help in recruiting and testing participants.

REFERENCES

Acuña-Ferreira, V. (2002). Gendered emotive displays in complaint discourse. *Estud. Sociolingüística* 3, 139–172.

Alicke, M. D., Braun, J. C., Glor, J. E., Klotz, M. L., Magee, J., Sederholm, H., et al. (1992). Complaining behavior in social interactions. *Pers. Soc. Psychol. Bull.* 18, 286–295.

Arndt, H., and Janney, R. W. (1991). Verbal, prosodic, and kinesic emotive contrasts in speech. *J. Pragmat.* 15, 521–549. doi: 10.1016/0378-2166(91)90110-J

Auger, J., and Valdman, A. (1999). Letting french students hear the diverse voices of francophony. *Mod. Lang. J.* 83, 403–412. doi: 10.1111/0026-7902.00030

Aziz-Zadeh, L., Sheng, T., Gheytanchi, A., and Aleman, A. (2010). Common premotor regions for the perception and production of prosody and correlations with empathy and prosodic ability. *PLoS One* 5:e8759. doi: 10.1371/journal.pone.0008759

Bates, D., Machler, M., Bolker, B. M., and Walker, S. C. (2015). Fitting linear mixed-effects models using Lme4. *J. Stat. Softw.* 67, 1–48. doi: 10.18637/jss.v067.i01

Boersma, P., and van Heuven, V. (2001). Speak and UnSpeak with praat. *Glot Int.* 5, 341–347.

Boxer, D. (1993). Social distance and speech behavior: the case of indirect complaints. *J. Pragmat.* 19, 103–125. doi: 10.1016/0378-2166(93)90084-3

Brunswik, E. (1956). *Perception and the Representative Design of Psychological Experiments*. Berkeley, CA: Univ of California Press.

Caballero, J. A., Vergis, N., Jiang, X., and Pell, M. D. (2018). The sound of Im/Politeness. *Speech Commun.* 102, 39–53. doi: 10.1016/j.specom.2018.06.004

Caffi, C., and Janney, R. W. (1994). Toward a pragmatics of emotive communication. *J. Pragmat.* 22, 325–373. doi: 10.1016/0378-2166(94)90115-5

Contreras-Huerta, L. S., Hielscher, E., Sherwell, C. S., Rens, N., and Cunnington, R. (2014). Intergroup relationships do not reduce racial bias in empathic neural responses to pain. *Neuropsychologia* 64, 263–270. doi: 10.1016/j.neuropsychologia.2014.09.045

Culpeper, J., and Terkourafi, M. (2017). "Pragmatic approaches (Im)politeness," in *The Palgrave Handbook of Linguistic (Im)Politeness*, eds J. Culpeper, M. Haugh, and D. Kádár (London: Palgrave Macmillan), 11–39. doi: 10.1057/978-1-137-37508-7_2

de Vignemont, F., and Singer, T. (2006). The empathic brain: how, when and why? *Trends Cogn. Sci.* 10, 435–441. doi: 10.1016/j.tics.2006.08.008

Drew, P. (1998). Complaints about transgressions and misconduct. *Res. Lang. Soc. Interact.* 31, 295–325. doi: 10.1080/08351813.1998.9683595

Drew, P., and Walker, T. (2009). Going too far: complaining, escalating and disaffiliation. *J. Pragmat.* 41, 2400–2414. doi: 10.1016/j.pragma.2008.09.046

Edwards, D. (2005). Moaning, whinging and laughing: the subjective side of complaints. *Discourse Stud.* 7, 5–29. doi: 10.1177/1461445605048765

Ekman, P., Friesen, W. V., O'Sullivan, M., Chan, A., Diacoyanni-Tarlatzis, I., Heider, K., et al. (1987). Universals and cultural differences in the judgments of facial expressions of emotion. *J. Pers. Soc. Psychol.* 53, 712–717. doi: 10.1037/0022-3514.53.4.712

Elfenbein, H. A., and Ambady, N. (2002). On the universality and cultural specificity of emotion recognition: a meta-analysis. *Psychol. Bull.* 128, 203–235. doi: 10.1037/0033-2909.128.2.203

Eyben, F., Scherer, K. R., Schuller, W., Sundberg, J., Andr, E., Busso, C., et al. (2016). The Geneva minimalistic acoustic parameter set (GeMAPS) for voice research and affective computing. *IEEE Trans. Affect. Comput.* 7, 190–202. doi: 10.1109/TAFFC.2015.2457417

Eyben, F., Wöllmer, M., and Schuller, B. (2010). "OpenSMILE – the munich versatile and fast open-source audio feature extractor," in *MM'10 – Proceedings of the ACM Multimedia 2010 International Conference* (New York, NY: ACM Press), 1459–1462. doi: 10.1145/1873951.1874246

Fiske, S. T., Cuddy, A. J. C., Glick, P., and Xu, J. (2002). A model of (Often Mixed) stereotype content: competence and warmth respectively follow from perceived status and competition. *J. Pers. Soc. Psychol.* 82, 878–902. doi: 10.1037//0022-3514.82.6.878

Floccia, C., Goslin, J., Girard, F., and Konopczynski, G. (2006). Does a regional accent perturb speech processing? *J. Exp. Psychol. Hum. Percept. Perform.* 32, 1276–1293. doi: 10.1037/0096-1523.32.5.1276

Foster, E. D., and Deardorff, A. (2017). Open science framework (OSF). *J. Med. Libr. Assoc.* 105, 203–206. doi: 10.5195/JMLA.2017.88

Frick, R. W. (1985). Communicating emotion: the role of prosodic features. *Psychol. Bull.* 97, 412–429. doi: 10.1037/0033-2909.97.3.412

Gelinas-Chebat, C., and Chebat, J.-C. (1992). Effects of two voice characteristics on the attitudes toward advertising messages. *J. Soc. Psychol.* 132, 447–459. doi: 10.1080/00224545.1992.9924724

Gilet, A.-L., Mella, N., Studer, J., and Grühn, D. (2013). Assessing dispositional empathy in adults: a French validation of the interpersonal reactivity index (IRI). *Can. J. Behav. Sci.* 45, 42–48. doi: 10.1037/a0030425

Gill, M. M. (1994). Accent and stereotypes: their effect on perceptions of teachers and lecture comprehension. *J. Appl. Commun. Res.* 22, 348–361. doi: 10.1080/00909889409365409

Grandjean, D., Bänziger, T., and Scherer, K. R. (2006). Intonation as an interface between language and affect. *Prog. Brain Res.* 156, 235–247. doi: 10.1016/S0079-6123(06)56012-1

Greenwald, A. G., McGhee, D. E., and Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: the implicit association test. *J. Pers. Soc. Psychol.* 74, 1464–1480. doi: 10.1037/0022-3514.74.6.1464

Grice, H. P. (1989). "Logic and conversation," in *Studies in the Way of Words*, ed. P. Grice (Cambridge, MA: Harvard University Press), 1–13. doi: 10.1093/acprof:oso/9780195133004.003.0001

Gussenhoven, C. (2004). *The Phonology of Tone and Intonation*. Cambridge: Cambridge University Press, doi: 10.1017/CBO9780511616983

Heblich, S., Lameli, A., and Riener, G. (2015). The effect of perceived regional accents on individual economic behavior: a lab experiment on linguistic performance, cognitive ratings and economic decisions. *PLoS One* 10:e0124732. doi: 10.1371/journal.pone.0113475

Hellbernd, N., and Sammler, D. (2016). Prosody conveys Speaker's intentions: acoustic cues for speech act perception. *J. Mem. Lang.* 88, 70–86. doi: 10.1016/j.jml.2016.01.001

Jiang, X., Gossack-Keenan, K., and Pell, M. D. (2020). To believe or not to believe? How voice and accent information in speech alter listener impressions of trust. *Q. J. Exp. Psychol.* 73, 55–79. doi: 10.1177/1747021819865833

Jiang, X., and Pell, M. D. (2017). The sound of confidence and doubt. *Speech Commun.* 88, 106–126. doi: 10.1016/j.specom.2017.01.011

Jiang, X., Sanford, R., and Pell, M. D. (2018). Neural architecture underlying person perception from in-group and out-group voices. *NeuroImage* 181, 582–597. doi: 10.1016/J.NEUROIMAGE.2018.07.042

Judd, C. M., Westfall, J., and Kenny, D. A. (2016). Experiments with more than one random factor: designs, analytic models, and statistical power. *Annu. Rev. Psychol.* 68, 601–625. doi: 10.1146/annurev-psych-122414-033702

Juslin, P. N. (2013). "Vocal affect expression: problems and promises," in *Evolution of Emotional Communication*, eds E. Altenmüller, S. Schmidt, and E. Zimmermann (Oxford: Oxford University Press), 252–274. doi: 10.1093/acprof:oso/9780199583560.003.0016

Juslin, P. N., and Laukka, P. (2003). Communication of emotions in vocal expression and music performance: different channels, same code? *Psychol. Bull.* 129, 770–814. doi: 10.1037/0033-2909.129.5.770

Kanske, P., Böckler, A., Trautwein, F. M., and Singer, T. (2015). Dissecting the social brain: introducing the EmpaToM to reveal distinct neural networks and brain-behavior relations for empathy and theory of mind. *NeuroImage* 122, 6–19. doi: 10.1016/j.neuroimage.2015.07.082

Katriel, T. (2013). "Griping' as a verbal ritual in some israeli discourse," in *Cultural Communication and Intercultural Contact*, ed. D. Carbaugh (New York, NY: Psychology Press), doi: 10.4324/9780203812259

Kircher, R. (2012). How pluricentric is the french language? An investigation of attitudes towards quebec french compared to European French. *J. French Lang. Stud.* 22, 345–370. doi: 10.1017/S0959269512000014

Kowalski, R. M. (2002). Whining, griping, and complaining: positivity in the negativity. *J. Clin. Psychol.* 58, 1023–1035. doi: 10.1002/jclp.10095

Kraft, B., and Geluykens, R. (2002). Complaining in French L1 and L2: a cross-linguistic investigation. *Eurosla Yearb.* 2, 227–242. doi: 10.1075/eurosla.2.14kra

Kuiper, L. (2005). Perception is reality: parisian and provencal perceptions of regional varieties of French1. *J. Sociolinguistics* 9, 28–52. doi: 10.1111/j.1360-6441.2005.00280.x

Kuznetsova, A., Brockhoff, P. B., and Christensen, R. H. B. (2017). LmerTest package: tests in linear mixed effects models. *J. Stat. Softw.* 82, 1–26. doi: 10.18637/jss.v082.i13

Laforest, M. (2002). Scenes of family life: complaining in everyday conversation. *J. Pragmat.* 34, 1595–1620. doi: 10.1016/S0378-2166(02)00077-2

Lang, S., Yu, T., Markl, A., Müller, F., and Kotchoubey, B. (2011). "Hearing others' pain: neural activity related to empathy. *Cogn. Affect. Behav. Neurosci.* 11, 386–395. doi: 10.3758/s13415-011-0035-0

Latoszek, B. B. v, De Bodt, M., Gerrits, E., and Maryn, Y. (2018). The exploration of an objective model for roughness with several acoustic markers. *J. Voice* 32, 149–161. doi: 10.1016/j.jvoice.2017.04.017

Laukka, P., and Elfenbein, H. A. (2020). Cross-cultural emotion recognition and in-group advantage in vocal expression: a meta-analysis. *Emot. Rev.* 1–9. doi: 10.1177/1754073919897295

Laukka, P., Elfenbein, H. A., Thingujam, N. S., Rockstuhl, T., Iraki, F. K., Chui, W., et al. (2016). The expression and recognition of emotions in the voice across five nations: a lens model analysis based on acoustic features. *J. Pers. Soc. Psychol.* 111, 686–705. doi: 10.1037/pspi0000066

Lautenbacher, S., Salinas-Ranneberg, M., Niebuhr, O., and Kunz, M. (2017). Phonetic characteristics of vocalizations during pain. *Pain Rep.* 2:e597. doi: 10.1097/PR9.0000000000000597

Lerner, A., Cohen, A., Avigal, M., Oshrat, Y., Bloch, A., and Zeilig, G. (2016). "Speech prosody as a biosignal for physical pain detection," in *Proceedings of the Conf Proc 8th Speech Prosody*, Boston, MA, 420–424. doi: 10.21437/SpeechProsody.2016-86

Lev-Ari, S., and Keysar, B. (2010). Why don't we believe non-native speakers? The influence of accent on credibility. *J. Exp. Soc. Psychol.* 46, 1093–1096. doi: 10.1016/j.jesp.2010.05.025

Liu, P., Rigoulot, S., Pell, M. D., Bernstein, L. E., and Tanaka, A. (2015). Cultural differences in on-line sensitivity to emotional voices: comparing east and west. *Front. Hum. Neurosci.* 9:311. doi: 10.3389/fnhum.2015.00311

Mac, D.-K., Aubergé, V., Rilliard, A., and Castelli, E. (2010). *Cross-Cultural Perception of Vietnamese Audio-Visual Prosodic Attitudes*. Chicago, IL: Speech Prosody.

Mauchand, M., Vergis, N., and Pell, M. D. (2018). Ironic Tones of Voices." In *Proceedings of the 9th International Conference on Speech Prosody 2018* (Poznan: Speech Prosody), 443–447. doi: 10.21437/SpeechProsody.2018-90

Mauchand, M., Vergis, N., and Pell, M. D. (2020). Irony, prosody, and social impressions of affective stance. *Discourse Process.* 57, 141–157. doi: 10.1080/0163853X.2019.1581588

Meconi, F., Doro, M., Lomoriello, A. S., Mastrella, G., and Sessa, P. (2018). Neural measures of the role of affective prosody in empathy for pain. *Sci. Rep.* 8:291. doi: 10.1038/s41598-017-18552-y

Ogden, R. (2010). *Prosody in Interaction*. Amsterdam: John Benjamins Pub. Co, 81–104.

Ohala, J. J. (1984). An ethological perspective on common cross-language utilization of Fo of Voice1. *Phonetica* 41, 1–16. doi: 10.1159/000261706

Ong, D. C., Zaki, J., and Goodman, N. D. (2018). Computational models of emotion inference in theory of mind: a review and roadmap. *Topics Cogn. Sci.* 11, 338–357. doi: 10.1111/tops.12371

Paulmann, S., and Kotz, S. A. (2008). An ERP investigation on the temporal dynamics of emotional prosody and emotional semantics in pseudo- and lexical-sentence context. *Brain Lang.* 105, 59–69. doi: 10.1016/j.bandl.2007.11.005

Pell, M. D. (2001). Influence of emotion and focus location on prosody in matched statements and questions. *J. Acoust. Soc. Am.* 109, 1668–1680. doi: 10.1121/1.1352088

Pell, M. D., Jaywant, A., Monetta, L., and Kotz, S. a (2011). Emotional speech processing: disentangling the effects of prosody and semantic cues. *Cogn. Emot.* 25, 834–853. doi: 10.1080/02699931.2010.516915

Pell, M. D., and Kotz, S. A. (2020). The next frontier: prosody research gets interpersonal. *Emot. Rev.*

Quadflieg, S., and Macrae, C. N. (2011). Stereotypes and stereotyping: what's the brain got to do with it? *Eur. Rev. Soc. Psychol.* 22, 215–273. doi: 10.1080/10463283.2011.627998

Raine, J., Pisanski, K., Simner, J., and Reby, D. (2019). Vocal communication of simulated pain. *Bioacoustics* 28, 404–426. doi: 10.1080/09524622.2018.1463295

Rao, R. (2013). Intonational variation in third party complaints in spanish. *J. Speech Sci.* 3, 141–168.

Regenbogen, C., Schneider, D. A., Finkelmeyer, A., Kohn, N., Derntl, B., Kellermann, T., et al. (2012). The differential contribution of facial expressions, prosody, and speech content to empathy. *Cogn. Emot.* 26, 995–1014. doi: 10.1080/02699931.2011.631296

Rigoulot, S., Jiang, X., Vergis, N., and Pell, M. D. (2020). Neurophysiological correlates of sexually evocative speech. *Biol. Psychol.* 154:107909. doi: 10.1016/j.biopsycho.2020.107909

Rodero, E. (2011). Intonation and emotion: influence of pitch levels and contour type on creating emotions. *J. Voice* 25, e25–e34. doi: 10.1016/J.JVOICE.2010.02.002

Sauter, D. A., Eisner, F., Ekman, P., and Scott, S. K. (2010). Cross-cultural recognition of basic emotions through nonverbal emotional vocalizations. *Proc. Natl. Acad. Sci. U.S.A.* 107, 2408–2412. doi: 10.1073/pnas.0908239106

Scarantino, A. (2017). How to do things with emotional expressions: the theory of affective pragmatics. *Psychol. Inquiry* 28, 165–185. doi: 10.1080/1047840X.2017.1328951

Scherer, K. R. (2003). Vocal communication of emotion: a review of research paradigms. *Speech Commun.* 40, 227–256. doi: 10.1016/S0167-6393(02)00084-5

Scherer, K. R., and Bänziger, T. (2004). "Emotional expression in prosody: a review and an agenda for future research," in *Speech Prosody 2004. Proceedings of the 2nd International Conference on Speech Prosody*, Nara, 359–366.

Scherer, K. R., Clark-Polner, E., and Mortillaro, M. (2011). In the eye of the beholder? Universality and cultural specificity in the expression and perception of emotion. *Int. J. Psychol.* 46, 401–435. doi: 10.1080/00207594.2011.626049

Searle, J. R. (1965). What is a speech act. *Perspectives in the Philosophy of Language: A Concise Anthology* 2000, 253–268.

Selting, M. (1994). Emphatic speech style?: with special focus on the prosodic signalling of heightened emotive involvement in conservation. *J. Pragmat.* 22, 375–408. doi: 10.1016/0378-2166(94)90116-3

Selting, M. (2010). Affectivity in conversational storytelling: an analysis of displays of anger or indignation in complaint stories. *Pragmatics* 20, 229–277. doi: 10.1075/prag.20.2.06sel

Soto, J. A., and Levenson, R. W. (2009). Emotion recognition across cultures: the influence of ethnicity on empathic accuracy and physiological linkage. *Emotion* 9, 874–884. doi: 10.1037/a0017399

Sumner, M., and Samuel, A. G. (2009). The effect of experience on the perception and representation of dialect variants. *J. Mem. Lang.* 60, 487–501. doi: 10.1016/j.jml.2009.01.001

Szczepek Reed, B. (2011). Beyond the particular: prosody and the coordination of actions. *Lang. Speech* 55, 13–34. doi: 10.1177/0023830911428871

Thierry, G., Paulmann, S., Jonczyk, R., Hatzidaki, A., Baus, C., and Costa, A. (2015). The way you say it, the way i feel it: emotional word processing in accented speech. *Front. Psychol.* 6:351. doi: 10.3389/fpsyg.2015.00351

Traunmüller, H., and Eriksson, A. (2000). Acoustic effects of variation in vocal effort by men, women, and children. *J. Acoust. Soc. Am.* 107, 3438–3451. doi: 10.1121/1.429414

Traverso, V. (2009). The dilemmas of third-party complaints in conversation between friends. *J. Pragmat.* 41, 2385–2399. doi: 10.1016/j.pragma.2008.09.047

Trenchs, M. (1994). Complaining in catalan, complaining in english: a comparative study of native and EFL speakers. *Rev. Esp. Lingüíst. Apl.* 10, 271–288.

Truesdale, D. M., and Pell, M. D. (2018). The sound of passion and indifference. *Speech Commun.* 99, 124–134. doi: 10.1016/J.SPECOM.2018.03.007

Van Kleef, G. A. (2009). How emotions regulate social life. *Curr. Dir. Psychol. Sci.* 18, 184–188. doi: 10.1111/j.1467-8721.2009.01633.x

Van Zant, A. B., and Berger, J. (2020). How the voice persuades. *J. Pers. Soc. Psychol.* 118, 661–682. doi: 10.1037/pspi0000193

Vanderweele, T. J., and Vansteelandt, S. (2014). Mediation analysis with multiple mediators. *Epidemiol. Method* 2, 95–115. doi: 10.1515/em-2012-0010

Wambacq, I. J. A., and Jerger, J. F. (2004). Processing of affective prosody and lexical-semantics in spoken utterances as differentiated by event-related potentials. *Cogn. Brain Res.* 20, 427–437. doi: 10.1016/J.COGBRAINRES.2004.03.015

Wichmann, A. (2000). *The Attitudinal Effects of Prosody, and How They Relate to Emotion*. Newcastle: ITRW on Speech and Emotion, 143–148.

Wichmann, A. (2002). "Attitudinal intonation and the inferential process. *Speech Prosody 2002*," in *Proceedings of the 1st International Conference on Speech Prosody*, (Aix-en-Provence: Speech Prosody), 11–16.

Winchatz, M. R. (2016). "Jammern [Whining] as a German way of speaking," in *The Handbook of Communication in Cross-Cultural Perspective*, ed. D. Carbaugh (New York, NY: Routledge), doi: 10.4324/9781315709321

Xu, X., Zuo, X., Wang, X., and Han, S. (2009). Do you feel my pain? racial group membership modulates empathic neural responses. *J. Neurosci.* 29, 8525–8529. doi: 10.1523/JNEUROSCI.2418-09.2009

Yuan, Y. C., Liao, W., and Bazarova, N. N. (2019). Judging expertise through communication styles in intercultural collaboration. *Manage. Commun. Q.* 33, 238–271. doi: 10.1177/0893318918824674

Zougkou, K., Weinstein, N., and Paulmann, S. (2017). ERP correlates of motivating voices: quality of motivation and time-course matters. *Soc. Cogn. Affect. Neurosci.* 12, 1687–1700. doi: 10.1093/scan/nsx064

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Mauchand and Pell. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

APPENDIX

TABLE A1 | Sentences constructed for the experiment with English translations. Only the bolded final words differed between the Pain and Control version of one sentence root.

Experimental Stimuli		***English translation***	
Pain	**Control**	***Pain***	***Control***
Ils ont tout fait sans **moi**	Ils ont tout fait sans **boire**	*They did everything without me*	*They did everything without drinking*
Ils sont partis sans **moi**	Ils sont partis sans **barque**	*They left without me*	*They left without a boat*
Ils m'ont demandé de **partir**	Ils m'ont demandé de **rester**	*They asked me to leave*	*They asked me to stay*
Ils refusé de venir chez **moi**	Ils ont refusé venir chez **Marc**	*They decided not to come to my place*	*They decided not to come to Marc's*
Ils ont refusé **m'inviter**	Ils ont refusé de **mélanger**	*They refused to invite me*	*They refused to stir*
Ils ont décidé de pas **m'inviter**	Ils ont décidé de pas **mélanger**	*They decided not to invite me*	*They decided not to stir*
Ils ont décidé de pas jouer avec **moi**	Ils ont décidé de pas jouer avec **Marc**	*They decided not to play with me*	*They decided not to play with Marc*
Ils ont décidé d'y aller sans **moi**	Ils ont décidé d'y aller sans **masques**	*The decided to go without me*	*They decided to go without masks*
Il m'a choisi en **dernier**	Il m'a choisi en **deuxiéme**	*He chose me last*	*He chose me in second*
Elle veut que personne **m'aime**	Elle veut que personne **marche**	*She wants no one to like me*	*She wants no one to walk*
Il me parle **jamais**	Il me parle jeudi	*He never talks to me*	*He talks to me thursday*
Elle a profité de **moi**	Elle a profité de **l'offre**	*She took advantage of me*	*She took advantage of the sale*
Il m'a fait **pleurer**	Il m'a fait **parler**	*He made me cry*	*He made me talk*
Elle m'a répondu **méchamment**	Elle m'a répondu **normalement**	*He answered me harshly*	*He answered me normally*
Elle a dit qu'elle m'aimait **pas**	Elle a dit qu'il m'aimait **bien**	*She said he didn't like me*	*She said he appreciated me*
Il veut vraiment pas de **moi**	Il veut vraiment pas de **masque**	*He really doesn't want me*	*He really doesn't want a mask*
Elle continue de **m'ignorer**	Elle continue de **mesurer**	*She keeps ignoring me*	*She keeps measuring*
Il a dit que j'étais **stupide**	Il a dit que j'étais **sorti**	*He said I was stupid*	*He said I went out*
Elle a dit que j'étais **sale**	Elle a dit que j'étais **jeune**	*She said I was dirty*	*She told me I was young*
Il a dit que j'étais **gros(se)**	Il a dit que j'étais **grand**	*He said I was fat*	*He said I was tall*
Elle pense que je suis **mauvaise**	Elle pense que je suis **bronzé**	*She thinks I am bad*	*She thinks I am tan*
Ils pensent que je suis **méchant(e)**	Ils pensent que je suis **belge**	*They think I am mean*	*They think I am belgian*
Elle pense que je suis **peureux (se)**	Elle pense que je suis **parti(e)**	*Shee thinks I am scared*	*She thinks I am gone*
Il pense que je suis **faible**	Il pense que je suis **fier**	*He thinks I am weak*	*He thinks I am proud*
Ils font des blagues sur **moi**	Ils font des blagues sur **Mars**	*They made jokes about me*	*They make jokes about Mars*
Ils font des blagues sur **mon poids**	Ils font des blagues sur **Montréal**	*They make jokes about my weight*	*They make jokes about Montreal*
Elle déteste **mon idée**	Elle déteste **mélanger**	*She hates my idea*	*She hates stirring*
Il m'a fait passer pour un **fou**	Il m'a fait passer pour un **frére**	*He made me look like madman*	*He made me look like a brother*
Elle m'a donné une **claque**	Elle m'a donné une **glace**	*Shee gave me a slap*	*She gave me an ice cream*
Il m'a donné un coup de **pied**	Il m'a donné un coup de **main**	*He gave me a kick*	*He gave me help*
Il m'empêche de **m'amuser**	Il m'empêche de **glisser**	*He prevents me from having fun*	*He rpevents me from slipping*
Elle m'empêche de **dormir**	Elle m'empêche de **tomber**	*She prevent me from sleeping*	*She prevents me from falling*
Il continue de me **mentir**	Il continue de me **montrer**	*He keeps lying to me*	*He keeps showing me*
Elle continue de **m'insulter**	Elle continue de **mélanger**	*She keeps insulting me*	*She keeps stirring*
Il a marché sur **mon chien**	Il a marché sur **mon chemin**	*He stepped on my dog*	*He stepped on my path*
Elle a marché sur **ma main**	Elle a marché sur **ma route**	*She stepped on my hand*	*She stepped on my road*
Il a frappé **ma jambe**	Il a frappé **ma balle**	*He kicked my leg*	*He kicked my ball*
Elle a pris **ma place**	Elle a pris **ma main**	*She took my spot*	*She took my hand*
Il m'a fait **tomber**	Il m'a fait **comprendre**	*He made me fall*	*He amde me understand*
Elle veut me faire **rater**	Elle veut me faire **rester**	*She wants to make me fail*	*She wants to make me stay*
Il essaye de **m'énerver**	Il essaye de **mesurer**	*He's trying to annoy me*	*He's trying to measure*

frontiers in Psychology

ORIGINAL RESEARCH
published: 05 February 2021
doi: 10.3389/fpsyg.2021.602023

Effects of Self-Expressive Brand and Susceptibility to Interpersonal Influence on Brand Addiction: Mediating Role of Brand Passion

Shizhen Bai[1], Yue Yin[1], Yubing Yu[2]*, Sheng Wei[1] and Rong Wu[1]*

[1] School of Management, Harbin University of Commerce, Harbin, China, [2] Logistics and E-commerce College, Zhejiang Wanli University, Ningbo, China

Although the concept of the consumer–brand relationship has undergone rapid change over the past two decades, the issue of brand addiction is still generally neglected in the literature. Based on social identity theory, the research develops a conceptual model of the influence of self-expressive brands (SEBs) and susceptibility to interpersonal influence (SUSCEP) on brand addiction. The results of this research demonstrate both separate and joint effects of SEBs and SUSCEP on brand addiction. In addition, harmonious brand passion and obsessive brand passion positively mediate the relationships among SEB, SUSCEP, and brand addiction. The research explores the formation mechanism of brand addiction from a new perspective and has important practical implications for brand marketers concerned with finding the most effective means to enhance the consumer–brand relationship.

Keywords: susceptibility to interpersonal influence, self-expressive brand, consumer-brand relationship, brand passion, brand addiction

OPEN ACCESS

Edited by:
Francesca D'Errico, University of Bari Aldo Moro, Italy

Reviewed by:
Maria Luisa Giancaspro, University of Bari Aldo Moro, Italy
Concetta Papapicco, University of Bari Aldo Moro, Italy

***Correspondence:**
Yue Yin
18724622417@163.com
Yubing Yu
yubing@zwu.edu.cn

Specialty section:
This article was submitted to Personality and Social Psychology, a section of the journal Frontiers in Psychology

Received: *02 September 2020*
Accepted: *13 January 2021*
Published: *05 February 2021*

Citation:
Bai S, Yin Y, Yu Y, Wei S and Wu R (2021) Effects of Self-Expressive Brand and Susceptibility to Interpersonal Influence on Brand Addiction: Mediating Role of Brand Passion. Front. Psychol. 12:602023. doi: 10.3389/fpsyg.2021.602023

INTRODUCTION

The concept of the customer–brand relationship initially focused on general attitudinal tendencies like commitment and loyalty, but researchers have expanded the concept to examine more intense forms of brand connection, such as brand love (Batra et al., 2012; Bagozzi et al., 2017), brand attachment (Jahn et al., 2012; Dolbec and Chebat, 2013), and, most recently, brand addiction (Cui et al., 2018; Mrad et al., 2020). The literature on brand addiction is still relatively nascent, and researchers are mainly concerned with its operationalization and conceptualization (e.g., Mrad and Cui, 2017, 2020). Because brand addiction represents the most intense or closest level of relation with brands, it is becoming a popular topic of study in brand marketing. However, given that brand addiction can lead to positive or negative outcomes (e.g., Mrad and Cui, 2017; Mrad et al., 2020), there is a strong impetus to understand its antecedents.

Swimberghe et al. (2014) proposed two broad categories of antecedents to an individual's behavior: the self and the influence of others. There is limited empirical research that takes both categories into consideration. Some studies have approached the issue of brand relationships from the self-identity perspective; these include Wallace et al. (2014), who explored the association between SEBs and brand love; Lee and Workman (2015), who attempted to identify the role of self-expressive brands in young consumers' brand relationships. Other research has highlighted the

influence of others by discussing the role that consumption plays in bringing people together—i.e., the 'linking value' of brand consumption (Cova, 1997). So far, there is no research that can fully prove how brand addiction is separately and jointly influenced by factors related to these two categories (i.e., the self and the influence of others). Identifying these factors is expected to help brand managers choose more appropriate ways to attract consumers and facilitate close and intense brand relationships. We try to highlight on the issue by exploring the effects of both the self and the influence of others on brand addiction.

Mrad and Cui (2017, 2020) developed a definition of brand addiction and began to explore its antecedent variables from different perspectives. At present, the factors known to influence brand addiction mainly relate to attitude (e.g., brand trust, brand attachment), perceived value, individual traits (e.g., perfectionism, self-enhancement), interpersonal influences, transcendent consumer experiences, and brand characteristics. Following Swimberghe, we consider antecedents related to customers' affective connections with brands (e.g., brand self-expression and brand passion), as these closely reflect customer commitment and have clear implications for brand addiction (Fournier, 1998; Albert et al., 2013). In addition, susceptibility to interpersonal influence can affect customers' brand choices (Ruane and Wallace, 2015). Indeed, the research has shown that in shaping customer attitude and behavior, the susceptibility to interpersonal influence has a role that cannot be ignored, and the greater the impact of interpersonal, the more likely consumers are to have strong feelings toward brands (Swimberghe et al., 2014).

Drawing on the insights above, our research explores how customers' affective connections with brands (i.e., brand self-expression and brand passion) and the influence of others affect brand addition. In a competitive landscape characterized by a large number of similar brands and products (Hwang and Kandampully, 2012), marketing strategies focused on price reductions and loyalty plans may be insufficient. A more promising strategy may be to help customers form intense brand connections, as strong consumer–brand relationships are believed to be the driving force in creating more sustainable brands (Park et al., 2006). Therefore, our research explores how customers' self-expression through brands and susceptibility to interpersonal influence affect their brand relationships and addictive behaviors. By examining the roles that the self and the influence of others play in this process, the results may help to deepen the understanding of brand addiction and have practical implications for brand marketers seeking efficient ways to enhance the consumer–brand relationship.

The rest of the study is organized as follows. In the second part, conceptual framework and theoretical background are presented. The third part focuses of hypotheses development. The fourth part describes the methodology, including the measures, data collection procedure, and sample. The fifth part estimates and assesses the structural model, evaluates the competitive models, and tests the mediating effects. The sixth part presents the results and primary conclusions, which includes implications for the literature and practitioners, limitations, future research direction of this research topic.

CONCEPTUAL FRAMEWORK AND THEORETICAL BACKGROUND

Social identity theory is considered appropriate for exploring the psychological mechanism of brand addiction, since identification has significant contribution to maintain the consumer–brand relationship (Lam et al., 2010). The social identity theory was developed by Tajfel and Turner who defined it "as that a part of an individual's self-concept which comes from the knowledge of his membership of social groups as well as the value and emotional significance attached to the membership" (Tajfel and Turner, 1979). The theory is usually associated with the studies of self-concept, and explains consumer behavior based on the interaction between self and society (Lam et al., 2010). On the one hand, the theory explains why consumers arouse strong feeling toward certain brands: consumers are eager to express their identities and differentiate themselves through certain brands. On the other hand, certain brands are consumed with the goal of gaining positive social recognition from groups to which customers hope to belong (Chernav et al., 2011; Sreejesh et al., 2016).

In addition, this study takes into account the psycho-social signals of "influence," which could constitute a theoretical basis of explanation of the binomial self/others (Poggi and D'Errico, 2012; Papapicco and Mininni, 2020). As Bearden and co-workers suggest, there is no adequate understanding of consumer behavior without considering the impacts of interpersonal influence on the individual's attitude, value, or consumption decision. Influence can be goal-directed, and occasionally individuals influence each other inadvertently (Poggi and D'Errico, 2012). The reason why consumers buy brand products is not only to meet the expectations of others, but also to send a signal to the consumer groups that they want to be recognized (Chernav et al., 2011).

This study explores the formation mechanism of brand addiction from the perspectives of the self and society. Based on previous research (Swimberghe et al., 2014), the constructs of self-expressive brands (SEBs) and susceptibility to interpersonal influence (SUSCEP) are used to reflect these two aspects. In addition, the shaping of the individual or social self may be autonomous or controlled in nature, and the resulting brand passion may be harmonious or obsessive. Considering the above, we investigate the mediating roles of harmonious brand passion (HBP) and obsessive brand passion (OBP) in the relationships among SEB, SUSCEP, and brand addiction to elucidate the process by which SEB and SUSCEP translate into brand addiction. The conceptual model guiding this research is outlined in **Figure 1**. The hypotheses are developed in the following sections.

RESEARCH HYPOTHESES

Self-Expressive Brands

An important part of the consumer–brand relationship is the idea of SEBs (Lee and Workman, 2015), which are defined as "the consumer's perception of the degree to which the specific brand

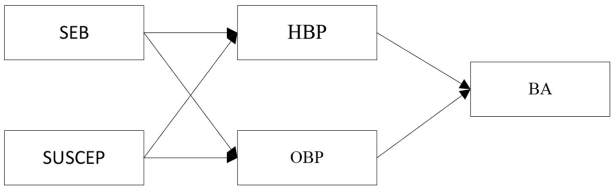


FIGURE 1 | Conceptual model. SEB, Self-expressive brand; SUSCEP, Susceptibility to interpersonal influence; HBP, Harmonious brand passion; OBP, obsessive brand passion; BA, Brand addiction.

enhances one's social self and/or reflects one's inner self" (Carroll and Ahuvia, 2006). To be specific, focus on personal self-concept, and the better ways to express themselves by brands (Loureiro et al., 2010). It is not difficult to find that the brand of self-expression is a symbol of personal achievement and a necessary carrier of social integration, which helps consumers better express their personal needs (Ruane and Wallace, 2015).Because brands that are more self-expressive tend to be more loved, managers can elicit more intense emotional responses from consumers by enhancing this aspect of their offerings. In addition, as consumers feel more love toward brands which are able to help them to shape their identity and self-expressiveness, it follows that SEBs should be one of the motivators of brand addiction (Mrad et al., 2020). A brand's self-expressive benefits can facilitate addiction, and a consumer–brand relationship can be maintained by enhancing the potential for self-expression (Ruane and Wallace, 2015).

The social identity theory suggests that individuals hope to express their own identities and differentiate themselves through certain brands. The existence of these brands provide a good way for individuals to achieve self-expression (Aaker, 2009), and their advantages also stimulate consumers' purchasing behavior. Combined with past research, it can be seen that customers who consume self-expression brands are more dependent on these brands than other customers (Ruane and Wallace, 2015). This conclusion is consistent with the argument of Liu et al. (2012). In short, when consumers think that the brand and their perceived self-image are the same, their trust and loyalty to the brand will increase. Most customers prefer to choose brands that match their own image or can symbolize their personal identity, and such brands are expected to elicit stronger emotional responses from consumers. In accordance with Aaker (2009), self-expressive brands can deepen the relationships between brands and consumers to a certain extent, and have also established that SEBs are an important factor affecting brand passion (Bauer et al., 2007).

Reviewing the literature related to marketing, it is found that there are many academic studies on brand passion (Swimberghe et al., 2014; Das et al., 2019; Mukherjee, 2019; Pourazad et al., 2019), and researchers have made corresponding definitions based on different backgrounds. The concepts have already accepted the explanation of brand passion by Swimberghe et al. (2014). They believe it reflects consumer's feeling for brands, which he/she values, blends into his/her identity, and invests resources. The definition assumes that there are close theoretical relationships between brand passion and self-identity, which also proves that the perspective of consumers' brand passion is consistent with the previous articles (Albert et al., 2013). There are two types of brand passion: harmonious and obsessive (Swimberghe et al., 2014). The difference lies in how a brand is internalized in a person's identity. When a brand projects an individual's identity through self-directed internalization, HBP gradually forms (Swimberghe et al., 2014). When a brand shows his/her identity through controlled internalization, OBP may result. The perception that a brand may enhance a consumer's identity generates positive emotions, leading to brand passion. Therefore, we propose:

H1. The level of SEB has a positive impact on HBP.

H2. The level of SEB has a positive impact on OBP.

Susceptibility to Interpersonal Influence

One of the most significant factors that determine an individual's behavior is the impact of others (Swimberghe et al., 2014). When consumers' behavior can be observed, these effects will emerge, and consumers' interpersonal influence will be directly reflected in their purchase behavior (Goudge et al., 2017). According to Bearden et al. (1989), SUSCEP is reflective of personality characteristics, which not only embodies certain need of a consumer to improve his/her own image through purchasing brand products, but also shows the willingness to comply with other people's expectations of purchasing decisions. Consumers who have a high SUSCEP are expected to prefer brands that manifest desirable traits for others to see Lertwannawit and Mandhachitara (2012), and have a stronger emotional response to the brands that symbolize their personal images (Astakhova et al., 2017). Mrad et al. (2020) showed that addiction to certain fast-fashion and luxury brands was associated with consequences related to interpersonal relations and financial issues.

When people consciously adjust their own deeds and beliefs to achieve consistency with others, the susceptibility is fairly obvious (Nabi et al., 2019). According to the psycho-social signals of "influence", the reason why consumers buy brand products is not only to meet the expectations of others, but also to send a signal to the consumer groups that they want to be recognized (Chernav et al., 2011; Sreejesh et al., 2016). Thus, a considerable number of consumers will show their group membership by purchasing some brand products. Individuals with higher susceptibility are also more willing to use interpersonal search to obtain the information they want to know (Mourali et al., 2005). As long as their information needs are met, the information will be regarded as real and can be used to meet the expectations of others (Kiani and Laroche, 2019). The consumers that desire to be recognized in this manner also expects to establish a good connection with society (Lertwannawit and Mandhachitara, 2012). Inner satisfaction is conducive to the autonomous internalization process of consumer–brand identification and the generation of HBP.

As far as consumers are concerned, the others' influence is mainly from colleagues, family members, and friends. The disapproval of their specific purchase behavior by these reference groups will bring great psychological pressure to

consumers and result in their blind pursuit of psychological convergence (Sharma and Klein, 2020). In this case, consumers buy brands because of interpersonal pressure, resulting in a brand relationship that is not controlled by the individual. This can lead to OBP, which originates from the controlled internalization process of consumer–brand identification (Das et al., 2019). Combined with existing research data, we can see that the impact of SUSCEP on OBP has been verified (Swimberghe et al., 2014), which also shows a consumer who is highly sensitive to interpersonal influences will be apt to have an obsessive passion for brands. Therefore, we propose:

H3. The level of SUSCEP has a positive impact on HBP.

H4. The level of SUSCEP has a positive impact on OBP.

The Mediating Role of HBP and OBP

The relationships between consumers and brands are usually closely related to the personality of consumers and the ways in which their interpersonal relationships form. When consumers have a high degree of recognition and trust for a certain brand, their emotion will be very positive. Social identity theory has been considered appropriate for understanding customers' psychological mechanisms (Lam et al., 2010), and the recognition effect can supplement other research topics, such as customer brand relationship (Swimberghe et al., 2014). The theory holds that consumers shape their private or social selves to define their self-concepts. Brands contribute to the private/social self construction on account of consumers regarding the brand as a medium to gain identity, which can be recognized by customers (Lam et al., 2010). The shaping of the individual and social roles may be autonomous or controlled, generating two forms of brand passion, i.e., HBP or OBP (Vallerand et al., 2003). Both of them effectively prove how a brand internalizes a person's identity.

Under the influence of various schemas, brand passion can realize the close relationship between brand and customers, and customers' demand for identity will last for a long time (Schmalz and Orth, 2012). HBP provides customers an internal motivation to lend importance to a brand, desire to get it without any contingencies or other influences (Albert et al., 2013). In the case of OBP, however, customers are compelled to purchase a brand due to interpersonal (social) or intra-personal (internal) pressures (Rauschnabel and Ahuvia, 2014), and this tends to lead to brand obsession and culminate in brand addiction (Fournier, 1998). Some authors consider the consumer-brand relationships are continuous points, and the relationships increase step by step (e.g., Fournier, 1998; Veloutsou and Moutinho, 2009). Thus, emotional passion toward a brand is a predictor of addiction.

Considering the continuum of consumer-brand relationships, we propose that brand passion is an element that ought to be taken into account. Based on the mediational hypothesis route, assuming that consumers regard a brand as one of their relationship partners, several brand clues that meet their own needs can help them to define themselves relatively accurately and to achieve consistency with others; this, in turn, triggers a strong emotional reaction in the form of brand passion, either harmonious or obsessive (Das et al., 2019). In other words, SEBs and SUSCEP may be considered important drivers of brand passion, which, in turn, leads to a higher tendency toward brand addiction. Accordingly, we hypothesize that SEBs and SUSCEP influence brand addiction through HBP and OBP. Thus, we propose:

H5. HBP mediates the effect of both SEB and SUSCEP on brand addiction.

H6. OBP mediates the effect of both SEB and SUSCEP on brand addiction.

RESEARCH METHODOLOGY

Ethics Approval Statements

The studies involving human participants were reviewed and approved by University of Commerce. All participants were consumers familiar with a certain brand finding through Harbin Mingyue Market Research Consulting Co., Ltd. using questionnaire. They verbally agreed to participate in this study and returned the questionnaire within a fixed time.

Data Collection and Sample

In the study, consumers with strong feelings about a particular brand were found through the snowball recommendation method. Respondents were asked to select favorite brands and fill in a questionnaire. Although snowball sampling is a non-random sampling process, this method significantly increases the likelihood of finding target consumers. There is evidence that its use can help researchers obtain high-quality and reliable data. A total of 800 questionnaires were distributed during the study period. 526 valid questionnaires were obtained after eliminating the questionnaires with missing value, contradictory answers.

TABLE 1 | Demographic information of respondents.

Criteria	Number	Percentage (%)
Sex		
Male	200	38.0%
Female	326	62.0%
Age		
18–25	261	49.6%
26–35	121	23.0%
36–45	114	21.7%
> 45	30	5.7%
Education		
Junior high school or lower	8	1.5%
High school/technical school	20	3.8%
Graduate	351	66.7%
Master's degree or higher	132	25.1%
Ph.D.	15	2.9%
Occupation		
Student	249	47.3%
Government or public institution employees	40	7.6%
Enterprise staff	121	23.0%
Self-employed	102	19.4%
Others	14	2.7%

The detailed list of the demographic information was showed in **Table 1**.

The sample consisted of consumers with strong feelings about a particular brand in China, ranging in ages from 18 to 55 (M = 34.88; SD = 8.237). Among the respondents, 62.0% were female, and 38.0% were male. About 47.3% of the respondents were students, 7.6% were employees of the government or public institutions, 23% were employees of enterprises, 19.4% were self-employed, and 2.7% were employed in other occupations.

Measures

Reliability analysis showed that Cronbach's alpha values were all above 0.70: SEB (0.88), SUSCEP (0.73), HBP (0.76), OBP (0.90), and brand addiction (0.88). Consequently, we preliminarily believes that the measurement values selected in the study have certain reliability.

The measurements of the constructs were validated based on previous research, which could be seen in **Table A1**. Regarding SEB, the eight-item scale of Carroll and Ahuvia (2006) was used; this included items such as "this brand reflects my personality" and "this brand is an extension of my inner self." The SUSCEP measures were taken from Lertwannawit and Mandhachitara (2012) as well as Ruane and Wallace (2015); these included items such as "I achieve a sense of belonging by purchasing the same products and brands that others purchase" and "I often identify with other people by purchasing the same products and brands they purchase." Regarding HBP and OBP, the measures of Swimberghe et al. (2014) were adopted; these included items such as "this activity allows me to live a variety of experiences" and "the urge is so strong that I can't help myself from doing this activity." Brand addiction was measured using the scales of Mrad and Cui (2017), with such statements as "I often find myself thinking about my favorite brand" and "I follow my favorite brand's news all the time." For the above measures, a 5-point Likert-type scale (1 = strongly disagree; 5 = strongly agree) was used.

DATA ANALYSIS AND RESULTS

Measurement Model

Items that correlated poorly with other items in all scales were detected by conducting a preliminary data analysis. As a result, one item of the original HBP scale and two items of the original brand addiction scale were deleted. A confirmatory factor analysis (CFA) was implemented for evaluating the measurement model's performance. The maximum likelihood (ML) estimation strategy was used to perform the estimation. This strategy assumes the multi-normality of the observed variables' distribution.

As advised by Kline (2017), kurtosis and skewness were evaluated to examine the departure from normality. Kurtosis in the data ranged from −1.19 to 1.8, and skewness ranged from −1.07 to 0.23. Thus, considering the thresholds (kurtosis < 20.0; skewness < 3.0) put forward by Kline (2017), the items met the assumptions of the ML estimation strategy. In spite of the results show that it is inconsistent with multiple normality, the application of ML estimation strategy can effectively solve this problem. Especially kurtosis needs to be tested because it has a large influence on the covariance and variance, which are the foundation of structural equation modeling. However, in accordance with the simulation research proposed by West et al. (1995), only the kurtosis value exceeds 7 is indicative of a serious departure from normality.

The model's fit indices acted well in terms of the acceptable thresholds in the core literature (Hair et al., 2006; Hwang and Kandampully, 2012). Although the chi-square (χ^2) test yielded a result of 708.466 with df = 199, which was statistically significant ($p < 0.01$), the remainder of the fit indices—i.e., the goodness of fit index (GFI; 0.88), the root mean square error approximation (RMSEA; 0.07), the comparative fit index (CFI; 0.91), the Tucker-Lewis index (TLI; 0.89), the incremental fit index (IFI; 0.91), and the normed fit index (NFI; 0.88)—implied that the measurement model was able to be accepted.

The standardized factor loadings were relatively large (every loading was over the 0.5 threshold) and statistically significant ($p < 0.01$). In addition, since R^2 values are all above the threshold value of 0.20, individual items are credible, and the convergence effectiveness of the above measures is also proved. By analyzing **Table 2**, we can clearly see that the comprehensive reliability of all the scales is greater than 0.70, indicating that the scales were internally consistent. For most of the structures in the model, the extracted average variance is not only greater than 50%, but also larger than the square of the correlation coefficients of the two potential variables. These results indicate the model's discriminant validity (see **Table 3**). Given the high correlation between HBP and brand addiction (r = 0.72), an alternative four-element model where brand addiction and HBP were loaded on a single element was also tested. This model yielded a worse fit (χ^2 = 848.270, df = 203, $p < 0.001$, RMSEA = 0.08, CFI = 0.88, TLI = 0.87, IFI = 0.88, NFI = .85, GFI = 0.86; $\Delta\chi^2$ = 139.804, Δdf = 4, $p < 0.001$), implying that HBP and brand addiction are distinct. In conclusion, the constructs present acceptable levels of discriminant validity, convergent validity, and reliability.

Structural Model

The model (shown in **Figure 1**) illustrates how both HBP and OBP mediate the impact of SEB and SUSCEP on brand addiction. It has been suggested that residuals linked to mediators be allowed to co-vary (Preacher and Hayes, 2008; Augusto and Torres, 2018). According to Gudmundsdottir et al. (2004), the residual correlation among mediators usually plays a substantive role. Thus, having the residuals of the mediator be steady at 0 is an unreasonable constraint that will cause the model to be misspecified. It was concluded that the mediators' errors were correlated by adhering to the recommendation applied in previous studies (e.g., Augusto and Torres, 2018).

The hypotheses and results of the structural model are presented in **Table 4**. The chi-square (χ^2) test yielded a value of 695.021 with df = 196 and $p < 0.01$. Moreover, the remaining fit indices implied a nice model fit (RMSEA = 0.07, CFI = 0.91, TLI = 0.89, IFI = 0.91, NFI = 0.88, GFI = 0.90). The results of the fit indices provide support to all the hypotheses. Meanwhile,

TABLE 2 | Confirmatory factor analysis results.

Construct	Dimension/Items	Stand. Loadings	*t*-value	CR	AVE
Self-expressive brand (SEB)	Inner self	0.82	–	0.79	0.66
	Social self	0.80	18.34		
Susceptibility to interpersonal influence (SUSCEP)	I achieve a sense of belonging by purchasing the same products and brands that others purchase.	0.70	–	0.73	0.47
	If I want to be like someone, I often try to buy the same brands that they buy.	0.61	10.22		
	I often identify with other people by purchasing the same products and brands they purchase.	0.74	11.63		
Harmonious brand passion (HBP)	This activity allows me to live a variety of experiences.	0.60	–	0.74	0.42
	The new things that I discover with this activity allow me to appreciate it even more.	0.68	11.60		
	This activity allows me to live memorable experiences.	0.67	11.08		
	This activity reflects the qualities I like about myself.	0.63	10.54		
Obsessive brand passion (OBP)	I cannot live without it.	0.77	–	0.90	0.65
	The urge is so strong, I can't help myself from doing this activity.	0.76	18.27		
	I am emotionally dependent on this activity.	0.81	18.94		
	I have a tough time controlling my need to do this activity.	0.87	20.15		
	I have almost an obsessive feeling for this activity.	0.82	18.81		
Brand addiction (BA)	I often find myself thinking about my favorite brand.	0.67	–	0.86	0.43
	I tend to give up some life activities and duties, such as the occupational, academic, and familial, to fulfill some activities related to my favorite brand.	0.60	12.28		
	I tend to allocate a certain portion of my monthly income to buy the products of my favorite brand.	0.61	12.53		
	I usually remember tenderly the previous experiences with my favorite brand.	0.68	13.80		
	I experience a state of impatience immediately before I can get hold of the products of my favorite brand.	0.61	12.48		
	I follow my favorite brand's news all the time.	0.67	13.51		
	I usually plan when the next purchase of my favorite brand will be.	0.70	14.05		
	I would invest my money in some way to support my favorite brand.	0.72	14.48		

the results indicate SEB has a stronger impact than SUSCEP on both HBP and OBP, while OBP and HBP have similar impacts on brand addiction.

TABLE 3 | Correlations and square root of AVE.

Construct	X_1	X_2	X_3	X_4	X_5
SEB(X_1)	***0.81***				
SUSCEP(X_2)	0.76	***0.69***			
HBP(X_3)	0.70	0.60	***0.65***		
OBP(X_4)	0.69	0.59	0.55	***0.81***	
BA(X_5)	0.77	0.65	0.72	0.72	***0.66***

The bold italic numbers in diagonal are the square roots of the average variances extracted.
SEB, Self-expressive brand; SUSCEP, Susceptibility to interpersonal influence; HBP, Harmonious brand passion; OBP, Obsessive brand passion; BA, Brand addiction.

Testing for Mediation Effects

We tested the mediating roles of OBP and HBP on the relationship between the dependent variable (brand addiction) and the independent variables (SUSCEP and SEB), three extra models were evaluated following the method used by James et al. (2006), Saenz et al. (2014), Augusto and Torres (2018), among others. The results of all four models are shown in **Table 5**. Model 1 is the base model. In Model 2, only the direct effects of SEB and SUSCEP on brand addiction were evaluated. Model 3 examined the direct impacts of SUSCEP and SEB on the dependent variable (brand addiction) and included the mediators. Finally, Model 4 corresponds to the base model, plus the direct effects of SEB and SUSCEP on brand addiction.

For the sake of verifying the existence of mediating effects, certain conditions must first be met. First, it is necessary for the independent variables to directly influence the mediators. Second, it is necessary for the mediators to directly affect the

TABLE 4 | Results of the structural model.

Path	Stand. coeff.	*t*-value	*p*-value	Hypotheses
SEB→HBP	0.68	8.74	***	H1(+): S
SEB→OBP	0.68	10.66	***	H2(+): S
SUSCEP→HBP	0.25	3.65	***	H3(+): S
SUSCEP→OBP	0.25	4.15	***	H4(+): S
HBP→BA	0.48	6.73	***	H5(+): S
OBP→BA	0.51	8.55	***	H6(+): S

Stand. coeff., standardized coefficient; SEB, Self-expressive brand; SUSCEP, Susceptibility to interpersonal influence; HBP, Harmonious brand passion; OBP, Obsessive brand passion; BA, Brand addiction; S, supported.
R^2: HBP: 0.52; OBP: 0.52; BA: 0.74.
$^{***}p < 0.001$.

TABLE 5 | Results of model estimation.

	Model 1, full mediation	Model 2	Model 3, non mediation	Model 4, partial mediation
SEB→HBP	0.68***		0.73***	0.67***
SEB→OBP	0.68***		0.75***	0.68***
SUSCEP→HBP	0.25***		0.19**	0.21**
SUSCEP→OBP	0.25***		0.21***	0.23***
SEB→BA	–	0.69***	0.82***	0.20*
SUSCEP→BA	–	0.27***	0.27***	0.15**
HBP→BA	0.48***		–	0.34***
OBP→BA	0.51***		–	0.40***
R^2				
HBP	0.52		0.57	0.49
OBP	0.52		0.61	0.51
BA	0.74	0.55	0.74	0.72

$^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

dependent variable. According to the evaluation result of model 1, the above two conditions are verified effectively. Third, it is necessary for the independent variable to directly affect the dependent variable with the absence of mediators. This condition was tested by an estimation of Model 2. The direct impacts of SUSCEP and SEB on brand addiction are important in the model. Finally, if the previous hypothesis turns out not to be supported after the mediators are included in the framework in the fourth step, or if their influence still exists but decreases, it indicates that there is partial mediation or full mediation. The results regarding mediating effects, which are shown in **Tables 5**, **6** indicate that Model 1 is a much nicer fit than Model 3 (non-mediation model) ($\Delta\chi^2 = 44.906$, $\Delta df = 0$, $p < 0.01$). Furthermore, compared with Model 1, Model 4 (a partial mediation model) is a nicer fit ($\Delta\chi^2 = 1.6$, $\Delta df = 2$, $p < 0.01$). In conclusion, given that the paths from SUSCEP and SEB to brand addiction still existed, but with reduced impacts, after including OBP and HBP, the partial mediation model was supported.

DISCUSSION

This research investigated the effects of self-expressive brands and susceptibility to interpersonal influence on brand addiction, considering harmonious brand passion and obsessive brand passion as mediators. Companies seek to create highly emotional consumer–brand relationships (Schmid and Huber, 2019), but satisfying consumer wants and needs may not be enough; Moreover, company should make accurate market positioning for its brands, so as to meet the demands of customers' life, and become an indispensible tool for customers to express themselves. Our findings indicate that SEBs have positive effects on both HBP (H1) and OBP (H2). These findings are consistent with prior studies that identified SEBs as a core contributor to customer–brand relationship enhancement (e.g., Swimberghe et al., 2014).

TABLE 6 | Results of model comparison.

	χ^2	*df*	Δdf	$\Delta\chi^2$	GFI	NFI	IFI	TLI	CFI	RMSEA
Model 1	685.021	196	–	–	0.90	0.88	0.91	0.89	0.91	0.07
Model 3	729.927	196	0	44.906	0.89	0.87	0.90	0.89	0.90	0.07
Model 4	683.426	194	2	1.6	0.90	0.88	0.91	0.89	0.91	0.07

This research also finds that SUSCEP has positive effects on OBP and HBP (H3 and H4). These results explain the importance of interpersonal effects, which should be one of the motivators of consumer's feeling for brands. The findings deepen understanding of the interplay among the constructs and indicate that SUSCEP might positively affect two forms of brand passion, meaning that a consumer who is sensitive to interpersonal influences can generate a harmonious or obsessive passion for brands.

At last, the results shows that the impacts of SUSCEP and SEBs on brand addiction are mediated by OBP and HBP. For brand managers, this implies a need to stimulate positive responses in terms of SUSCEP and SEBs, as this leads to brand passion and in turn has significant effects on brand addiction (H5 and H6). In line with Hatfield (1988), we support that passion relate to more intense emotional responses; emotional passion toward a brand is the antecedent of brand addiction. Moreover, brand managers should spend resources on a "passion branding" strategy and recognize the distinction between the two kinds of brand passion.

Theoretical Implications

First, we explored the effects of SEBs and SUSCEP on brand addiction through quantitative research methods. The literature on brand addiction focuses on its conceptualization, trying to distinguish the following concepts, such as compulsive purchase or other forms of consumer-brand relationship (e.g., Mrad and Cui, 2017, 2020). Nevertheless, most of the related studies are qualitative in nature. For instance, Cui et al. (2018) analyzed 11 salient features of brand addiction by using qualitative data from projective interviewing and focus groups, and Mrad et al. (2020) did in-depth interviews to examine the motives behind consumers' addictive behavior. The results extend current research by responding to the call of Mrad and Cui (2017) for a quantitative study to complement their findings.

Second, our research provides a new perspective for exploring the formation mechanism of brand addiction. In the light of social identity theory, the research finds SEBs, SUSCEP,

and brand passion play crucial roles in the formation of brand addiction. These results provide conceptual frameworks to further understand the internal mechanism of brand addiction. Thus, this study supports the distinct treatment of the relationships between consumers and brands based on data analysis.

Our findings also confirm that SEBs and SUSCEP are positively related to HBP and OBP. A consumer who perceives brands to enhance his or her identify may gradually develop harmonious or obsessive passion for those brands. As anticipated, and in line with Swimberghe et al. (2014), SEBs have positive impact on harmonious passion and obsessive passion. Meanwhile, considering that most consumers needs to identify with others or meet their expectations (Kurt et al., 2011), it is understandable that SUSCEP has positive effects on both HBP and OBP. This lends further support to the argument of Belk et al. (2003) that brand passion tends to be more prominent under interpersonal influence. Our findings show how the psychological effects generated by consumers' SUSCEP can promote HBP and OBP.

Third, after the analysis of the mediating effect of HBP and OBP, the research further reveals the influence path of SEBs and SUSCEP on consumer brand addiction, and the empirical results also effectively prove that the structural equation model proposed by the author is feasible. The study found that HBP and OBP have a certain mediating effect, which shows that HBP and OBP play key roles in the effectiveness of SEBS and SUSCEP in strengthening brand addiction. In view of the fact that the research about brand addiction is still at the first stage, on the one hand, the findings will help to encourage experts and scholars to continue to carry out special research on brand addiction, on the other hand, it can also significantly mobilize scholars' interest in the joint effect of brand addiction and other forms of consumer brand relationship, so as to expand the research scope of this subject.

In addition, by clarifying the formation mechanism of brand addiction, this research helps to explain its positive and negative outcomes. To date, some research on brand addiction has suggested that addictive relationships with brands are destructive and progressive (e.g., Fournier and Alvarez, 2013), but the latest evidence shows brand addiction may not cause detrimental effects, partly because of its positive correlation to self-esteem (Cui et al., 2018). Furthermore, Mrad et al. (2020) believes that brand addiction is like a double-edged sword, which may reduce or improve the happiness of consumers, revealing not all patterns of addictive behavior are just pathological. Our study finds that both HBP and OBP are positively associated with brand addiction. We believe our research supports the assertions of Mrad et al. (2020) about brand addiction. Indeed, the desire to get it and willing to put in the effort can generate inner satisfaction in consumers, which is conducive to the autonomous internalization process of consumer–brand identification and the generation of positive brand addiction. However, if the relationship with the brand is not controlled by the consumer, OBP can form, leading to the emergence of negative brand addiction. Our findings indicate that HBP and OBP appear to have similar effects on brand addiction.

Finally, this study addresses the call by Swimberghe et al. (2014) to explore the relationship between OBP and brand addiction, and it supplements prior studies (e.g., Mrad et al., 2020) on the consumer–brand relationship continuum by adding two key variables that impact on brand addiction—i.e., HBP and OBP. The results could help marketing scholars or clinical psychologists to deepen the understanding of addiction and find out the potential adverse effects.

Managerial Implications

The study has some implications for marketers, particularly those who want to establish a highly emotional consumer–brand relationship. Our findings can inform brand managers on what factors to prioritize to generate more intense consumer–brand connections, thus helping them to develop a scientific and efficient brand strategy, and point out the direction for improving the efficiency of customer relationship management. Marketers can use our results to classify consumers and segment their markets according to different levels of consumer–brand relationship intensity.

By exploring the key antecedents of brand addiction, the study can help marketers to clarify key roles of SEBs and SUSCEP. Managers should consider customers' desire for self-expression and their need to connect with others. When some customers show their personal identity or values, they often give priority to the brand that suits them instead of outstanding in the market. Brand managers can cater to these tendencies by providing a forum where customers can express themselves and satisfy their desire for self-confirmation but also form connections with other similar consumers and be affected by others' brand selections. Brand managers should ensure that members are able to freely interact in a friendly online community, which will enable marketers to examine perceptions and emotions toward their brands in real time.

Despite the significant contribution of SUSCEP, SEBs have a greater effect on OBP and HBP. This suggests that emphasizing self-expression in brand positioning is more important to enhance the consumer–brand relationship. Brand managers seeking to build more intense connections with their brands and engender passion for their products should focus on the self expression of the brand. For instance, marketing promotion may encourage consumers to take the brand as a breakthrough, and take the initiative to form close contact with it, so that their own needs can be expressed and transmitted. Furthermore, managers can seek to highlight the distinctive personality of a brand from multiple angles (such as design, brand name, advertising, brand endorser), allowing consumers to judge intuitively which brand suits them and expresses their identity and value. Companies can build stronger brands by identifying the aspects of customers' behaviors and lifestyles most relevant to their need for self-expression.

In addition, the mediating roles of OBP and HBP identified in this research underscore the need for a "passion branding" strategy. Albert et al. (2013) argued that if certain brands play important roles in the process of a consumer's identity acquisition, then his/her emotion toward the brands will be stronger. Additionally, brand managers can seek ways

for consumers to gain positive social recognition and self-worth through their brands, thereby promoting brand emotion. Activating these kinds of intense emotional responses is important to improve a brand's competitive position.

Empirical evidence attests to the nature and complexity of brand addiction. The views obtained can provide certain reference for brand managers to maintain the relationship between brands and consumers. A key question, however, is how brand managers should respond to the negative consequences of brand addiction. As addiction to brands may cause negative consequences at times, brand managers should adopt different marketing strategies. For example, for consumers who have the ability to balance the relationship between their own life and brand passion, brand managers can continuously try to deepen their relationship with brands, hence helping them to experience positive brand addiction. However, for consumers whose brand addiction becomes dominant and causes an imbalance with other areas of life, brand managers can design small packaging or enforce purchase limitations to reduce the psychological pressure.

Limitations and Future Research

The study still has many deficiencies, so it needs to be further improved in the future. First, the research payed attention to the fashion context, with the respondents from only one country. Therefore, the results of this study may not be directly applied to other situations. Future research can extend this investigation to other countries and investigate differences among brand categories. Second, the study was based on social identity theory. Future research can apply other theoretical models to explore variables related to brand addiction. Third, the scope of this research was limited to the drivers or antecedents of brand addiction, which were analyzed to understand the formation mechanism and motivations for brand addiction. In future research, the influence of consumer brand addiction on subsequent behaviors can be analyzed. Finally, although our research considered brand passion and responded to the call of Mrad and Cui (2017) to explore the role of brand addiction within an integrated framework of consumer–brand relationships, there is a need to consider other factors (e.g., compulsive buying and other types of addictive behaviors) to further deepen the understanding of brand addiction.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the University of Commerce. All participants were consumers familiar with a certain brand finding through the Harbin Mingyue Market Research Consulting Co., Ltd. using questionnaire. They verbally agreed to participate in this study and returned the questionnaire within a fixed time.

AUTHOR CONTRIBUTIONS

SB and YYi: conceptualization. YYi and SW: methodology. SW: software. SW and RW: formal analysis and investigation. RW: resources. YYi: writing-original draft preparation. YYi and YYu: writing-review and editing. SB and YYu: supervision. SB: project administration and funding acquisition. All authors contributed to the article and approved the submitted version.

FUNDING

This research was supported by the National Natural Science Foundation of China (71671054). Youth innovative talents project of Harbin University of Commerce (No. 2020CX16), Innovative research project for graduate students of Harbin University of Commerce (YJSCX2018-475HSD).

REFERENCES

Aaker, D. (2009). Beyond functional benefits. *Mark. News* 9:23.

Albert, N., Merunk, A. D., and Valette, F. P. (2013). Brand passion: antecedents and consequences. *J. Bus. Res.* 66, 904–909. doi: 10.1016/j.jbusres.2011.12.009

Astakhova, M., Swimberghe, K. R., and Wooldridge, B. R. (2017). Actual and ideal-self congruence and dual brand passion. *J. Consum. Mark.* 34, 664–672. doi: 10.1108/jcm-10-2016-1985

Augusto, M., and Torres, P. (2018). Effects of brand attitude and eWOM on consumers' willingness to pay in the banking industry: mediating role of consumer-brand identification and brand equity. *J. Retailing Consum. Serv.* 42, 1–10. doi: 10.1016/j.jretconser.2018.01.005

Bagozzi, R. P., Batra, R., and Ahuvia, A. (2017). Brand love: development and validation of a practical scale. *Mark. Lett.* 28, 1–14. doi: 10.1007/s11002-016-9406-1

Batra, R., Ahuvia, A., and Bagozzi, R. P. (2012). Brand love. *J. Mark.* 76, 1–16.

Bauer, H. H., Heinrich, D., and Marin, I. (2007). "How to create high emotional consumer-brand relationships? The causalities of brand passion," in *2007 Australian and New Zealand Marketing Academy Conference Proceedings*, 2189–2198.

Bearden, W. O., Netemeyer, R. G., and Teel, J. E. (1989). Measurement of consumer susceptibility to interpersonal influence. *J. Consum. Res.* 15, 473–481. doi: 10.1086/209186

Belk, R., Ger, G., and Askegaard, S. (2003). The fire of desire: A multiple inquiry into consumer passion. *J. Consum. Res.* 30, 326–351.

Carroll, B. A., and Ahuvia, A. C. (2006). Some antecedents and outcomes of brand love. *Mark. Lett.* 17, 79–89. doi: 10.1007/s11002-006-4219-2

Chernav, A., Hamilton, R., and Gal, D. (2011). Competing for consumer identity: limits to self-expression and the perils of lifestyle branding. *J. Mark.* 75, 66–82. doi: 10.1509/jmkg.75.3.66

Cova, B. (1997). Community and consumption: towards a definition of the 'linking value' of products and services. *Eur. J. Mark.* 31, 297–316. doi: 10.1108/03090569710162380

Cui, C. C., Mrad, M., and Hogg, M. K. (2018). Brand addiction: exploring the concept and its definition through an experiential lens. *J. Bus. Res.* 87, 118–127. doi: 10.1016/j.jbusres.2018.02.028

Das, G., Agarwal, J., Malhotra, N. K., and Varshneya, G. (2019). Does brand experience translate into brand commitment? A mediated-moderation model of brand passion and perceived brand ethicality. *J. Bus. Res.* 95, 479–490. doi: 10.1016/j.jbusres.2018.05.026

Dolbec, P. Y., and Chebat, J. C. (2013). The impact of a flagship vs. a brand store on brand attitude, brand attachment and brand equity. *J. Retailing* 89, 460–466. doi: 10.1016/j.jretai.2013.06.003

Fournier, S. (1998). Consumers and their brands: developing relationship theory in consumer research. *J. Consum. Res.* 24, 343–373. doi: 10.1086/209515

Fournier, S., and Alvarez, C. (2013). Relating badly to brands. *J. Consum. Psychol.* 23, 253–264. doi: 10.1016/j.jcps.2013.01.004

Goudge, D., Good, M. C., Hyman, M. R., and Aguirre, G. (2017). Modeling specialty store customers' buy/no-buy decisions. *Int. J. Retail Distrib. Manag.* 45, 1260–1276. doi: 10.1108/ijrdm-03-2017-0036

Gudmundsdottir, B. M. A., Beck, J. G., Coffey, S. F., Miller, L. B. S., and Palyo, S. A. (2004). Quality of life and post trauma symptomatology in motor vehicle accident survivors: the mediating effects of depression and anxiety. *Depress. Anxiety* 20, 187–189. doi: 10.1002/da.20037

Hair, J. F., Black, B., Babin, B., Anderson, R. E., and Tatham, R. L. (2006). *Multivariate Data Analysis*, 6th Edn. Upper Saddle River, NJ: Prentice Hall.

Hatfield, E. (1988). "Passionate and companionate love," in *The Psychology of Love*, eds R. J. Sternberg and M. L. Barnes (New Haven, CT: Yale University Press), 191–217.

Hwang, J., and Kandampully, J. (2012). The role of emotional aspects in younger consumer-brand relationships. *J. Prod. Brand Manag.* 21, 98–108. doi: 10.1108/10610421211215517

Jahn, S., Gaus, H., and Kiessling, T. (2012). Trust, commitment, and older women: exploring brand attachment differences in the elderly segment. *Psychol. Mark.* 29, 445–457. doi: 10.1002/mar.20533

James, L. R., Mulaik, S. A., and Brett, J. M. (2006). A tale of two methods. *Organ. Res. Methods* 169, 575–592. doi: 10.1007/s11229-008-9437-0

Kiani, I., and Laroche, M. (2019). From desire to help to taking action: effects of personal traits and social media on market mavens' diffusion of information. *Psychol. Mark.* 36, 1147–1161. doi: 10.1002/mar.21263

Kline, R. B. (2017). *Principles and Practice of Structural Equation Modeling*, 3rd Edn. New York: The Guildford Press.

Kurt, D., Inman, J. J., and Argo, J. J. (2011). The influence of friends on consumer spending: the role of agency-communion orientation and self-monitoring. *J. Mark. Res.* 48, 741–754. doi: 10.1509/jmkr.48.4.741

Lam, S. K., Ahearne, M., Hu, Y., and Schillewaert, N. (2010). Resistance to brand switching when a radically new brand is introduced: a social identity theory perspective. *J. Mark.* 74, 128–146. doi: 10.1509/jmkg.74.6.128

Lee, S. H., and Workman, J. E. (2015). Determinants of brand loyalty: self-construal, self-expressive brands, and brand attachment. *Int. J. Fashion Design Technol. Educ.* 8, 12–20. doi: 10.1080/17543266.2014.974691

Lertwannawit, A., and Mandhachitara, R. (2012). Interpersonal effects on fashion consciousness and status consumption moderated by materialism in metropolitan men. *J. Bus. Res.* 65, 1408–1416. doi: 10.1016/j.jbusres.2011.10.006

Liu, F., Li, J., Mizerski, D., and Soh, H. (2012). Self-congruity, brand attitude, and brand loyalty: a study on luxury brands. *Eur. J. Mark.* 46, 922–937. doi: 10.1108/03090561211230098

Loureiro, S. M. C., Ruediger, K. H., and Demetris, V. (2010). Brand emotional connection and loyalty. *J. Brand Manag.* 20, 13–27. doi: 10.1057/bm.2012.3

Mourali, M., Laroche, M., and Pons, F. (2005). Individualistic orientation and consumer susceptibility to interpersonal influence. *J. Serv. Mark.* 19, 164–173. doi: 10.1108/08876040510596849

Mrad, M., and Cui, C. C. (2017). Brand addiction: conceptualization and scale development. *Eur. J. Mark.* 51, 1938–1960. doi: 10.1108/ejm-10-2016-0571

Mrad, M., and Cui, C. C. (2020). Comorbidity of compulsive buying and brand addiction: an examination of two types of addictive consumption. *J. Bus. Res.* 113, 399–408. doi: 10.1016/j.jbusres.2019.09.023

Mrad, M., Majdalani, J. F., Cui, C., and Khansa, Z. E. (2020). Brand addiction in the contexts of luxury and fast-fashion brands. *J. Retailing Consum. Serv.* 55:102089. doi: 10.1016/j.jretconser.2020.102089

Mukherjee, K. (2019). Social media marketing and customers' passion for brands. *Mark. Intell. Plan.* 38, 509–522. doi: 10.1108/mip-10-2018-0440

Nabi, N., O'Cass, A., and Siahtiri, V. (2019). Status consumption in newly emerging countries: the influence of personality traits and the mediating role of motivation to consume conspicuously. *J. Retailing Consum. Serv.* 46, 173–178. doi: 10.1016/j.jretconser.2017.09.009

Papapicco, C., and Mininni, G. (2020). Impact memes: PhDs HuMor(e). *Multimedia Tools Appl.* 79, 35973–35994. . https://doi.org/10.1007/s11042-020-09166-0,

Park, C. W., Macinnis, D. J., and Priester, J. R. (2006). Beyond attitude: attachment and consumer behavior. *Seoul J. Bus.* 12, 3–36.

Poggi, I., and D'Errico, F. (2012). Social signals: a framework in terms of goals and beliefs. *Cogn. Process.* 13, 427–445. doi: 10.1007/s10339-012-0512-6

Pourazad, N., Stocchi, L., and Pare, V. (2019). The power of brand passion in sports apparel brands. *J. Product Brand Manage.* 29, 547–568. doi: 10.1108/jpbm-12-2018-2164

Preacher, K. J., and Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behav. Res. Methods* 40, 879–891. doi: 10.3758/brm.40.3.879

Rauschnabel, P. A., and Ahuvia, A. C. (2014). You're so lovable: anthropomorphism and brand love. *J. Brand Manage.* 21, 372–395. doi: 10.1057/bm.2014.14

Ruane, L., and Wallace, E. (2015). Brand tribalism and self-expressive brands: social influences and brand outcomes. *J. Product Brand Manage.* 24, 333–348. doi: 10.1108/jpbm-07-2014-0656

Saenz, M. J., Revilla, E., and Knoppen, D. (2014). Absorptive capacity in buyer–supplier relationships: empirical evidence of its mediating role. *J. Supply Chain Manage.* 50, 18–40. doi: 10.1111/jscm.12020

Schmalz, S., and Orth, U. R. (2012). Brand attachment and consumer emotional response to unethical firm behavior. *Psychol. Mark.* 29, 869–884. doi: 10.1002/mar.20570

Schmid, D., and Huber, F. (2019). Brand love: emotionality and development of its elements across the relationship lifecycle. *Psychol. Mark.* 36, 305–320. doi: 10.1002/mar.21180

Sharma, V. M., and Klein, A. (2020). Consumer perceived value, involvement, trust, susceptibility to interpersonal influence, and intention to participate in online group buying. *J. Retailing Consum. Serv.* 52:101946. doi: 10.1016/j.jretconser.2019.101946

Sreejesh, S., Sarkar, A., and Roy, S. (2016). Validating a scale to measure consumer's luxury brand aspiration. *J. Product Brand Manage.* 25, 465–478. doi: 10.1108/jpbm-06-2014-0647

Swimberghe, K. R., Astakhova, M., and Wooldridge, B. R. (2014). A new dualistic approach to brand passion: harmonious and obsessive. *J. Bus. Res.* 67, 2657–2665. doi: 10.1016/j.jbusres.2014.04.003

Tajfel, H., and Turner, J. (1979). An integrative theory of intergroup conflict. *Soc. Psychol. Intergroup Relat.* 33, 94–109.

Vallerand, R. J., Blanchard, C., Mageau, G. A., Koestner, R., Ratelle, C., Leonard, M., et al. (2003). Les passions de l'ame: on obsessive and harmonious passion. *J. Personal. Soc. Psychol.* 23, 756–767. doi: 10.1037/0022-3514.85.4.756

Veloutsou, C., and Moutinho, L. (2009). Brand relationships through brand reputation and brand tribalism. *J. Bus. Res.* 62, 314–322. doi: 10.1016/j.jbusres.2008.05.010

Wallace, E., Bull, I., and Chernatony, L. D. (2014). Consumer engagement with self-expressive brands: brand love and WOM outcomes. *J. Product Brand Manage.* 23, 33–42.

West, S. G., Finch, J. F., and Curran, P. J. (1995). "Structural equation models with nonnormal variables: problems and remedies," in *Structural Equation Modeling: Concepts, Issues, and Applications*, ed. R. H. Hoyle (Thousand Oaks, CA: Sage), 56–75.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Bai, Yin, Yu, Wei and Wu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

APPENDIX

TABLE A1 | Measures.

Construct	Items	Mean	SD
Inner self	This brand symbolizes the kind of person I really am inside.	3.83	0.97
	This brand reflects my personality.	3.93	0.94
	This brand is an extension of my inner self Source: Colliander and Dahlén (2011).	4.15	0.77
	This brand mirrors the real me.	3.79	0.99
Social self	This brand contributes to my image.	3.89	0.97
	This brand adds to a social "role" I play.	3.78	1.03
	This brand has a positive impact on what others think of me.	3.74	1.00
	This brand improves the way society views me.	3.48	1.18
Susceptibility to interpersonal influence	I achieve a sense of belonging by purchasing the same products and brands that others purchase.	3.69	1.05
	If I want to be like someone, I often try to buy the same brands that they buy.	3.09	1.25
	I often identify with other people by purchasing the same products and brands they purchase.	3.38	1.15
Harmonious brand passion	This activity allows me to live a variety of experiences.	3.93	0.84
	The new things that I discover with this activity allow me to appreciate it even more.	4.00	0.87
	This activity allows me to live memorable experiences.	3.72	1.00
	This activity reflects the qualities I like about myself.	4.04	0.82
	This activity is in harmony with the other activities in my life.	3.90	0.95
Obsessive brand passion	I cannot live without it.	3.32	1.25
	The urge is so strong. I can't help myself from doing this activity.	3.56	1.17
	I am emotionally dependent on this activity.	3.40	1.17
	I have a tough time controlling my need to do this activity.	3.25	1.27
	I have almost an obsessive feeling for this activity.	3.14	1.34
Brand addiction	I try very hard to get everything from my favorite brand.	2.99	1.39
	I often fail to control myself from purchasing products of my favorite brand.	3.07	1.37
	I often find myself thinking about my favorite brand.	3.73	1.08
	I tend to give up some life activities and duties, such as the occupational, academic, and familial, to fulfill some activities related to my favorite brand.	2.80	1.29
	I tend to allocate a certain portion of my monthly income to buy the products of my favorite brand.	3.44	1.13
	I usually remember tenderly the previous experiences with my favorite brand.	3.79	0.95
	I experience a state of impatience immediately before I can get hold of the products of my favorite brand.	3.78	1.09
	I follow my favorite brand's news all the time.	3.82	1.08
	I usually plan when the next purchase of my favorite brand will be.	3.88	1.08
	I would invest my money in some way to support my favorite brand.	3.78	1.14

frontiers in Communication

ORIGINAL RESEARCH
published: 18 March 2021
doi: 10.3389/fcomm.2021.617532

Sound-Symbolic Semantics of Written Japanese Vowels in a Paper-Based Survey Study

Misa Ando[1†], Xinyi Liu[2†], Yan Yan[2†], Yutao Yang[2†], Shushi Namba[3†], Kazuaki Abe[2] and Toshimune Kambara[2]*

[1]Program in Psychology, School of Education, Hiroshima University, Higashihiroshima, Japan, [2]Department of Psychology, Graduate School of Education, Hiroshima University, Higashihiroshima, Japan, [3]Psychological Process Team, BZP, Robotics Project, RIKEN, Kyoto, Japan

Verbal sounds can be associated with specific meanings, a phenomenon called sound symbolism. Previous findings of sound symbolism have shown that words including specific consonants or vowels or mouth shapes to pronounce specific vowels associate with specific and subjective physical and emotional evaluations. The purpose of this study was to examine whether each written vowel in a given language was individually associated with specific subjective evaluations. Six hundred and thirteen participants used five-point semantic differential scales for 10 features (size, closeness, thickness, width, weight, height, depth, affection, excitement and familiarity) to rate written Japanese vowels (*a*, *i*, *u*, *e*, and *o*). The results showed that the size, closeness, thickness and width of *a*, *u* and *o* were significantly higher than those of *i* and *e*, whereas the affection and familiarity of *a* were higher than the others. These results were consistent with previous findings in which vowels in sound-symbolic words have been associated with physical (i.e., size, closeness, thickness and width) and emotional (i.e., affection) evaluations. Our findings suggest that each written Japanese vowel itself, with its individual characteristics, could individually contribute to specific and subjective physical and emotional evaluations. These findings provide insights on how we could better use letters for communicative relationships among writers and readers.

Keywords: semantic differential scales, vowels, sound symbolism, japanese, physical evaluation, emotional evaluation

OPEN ACCESS

Edited by:
Oliver Niebuhr, University of Southern Denmark, Denmark

Reviewed by:
Niklas Erben Johansson, Lund University, Sweden
Christophe COUPE, The University of Hong Kong, Hong Kong

***Correspondence:**
Toshimune Kambara
tkambara@hiroshima-u.ac.jp

†These authors have contributed equally to this work

Specialty section:
This article was submitted to Language Sciences, a section of the journal Frontiers in Communication

Received: *15 October 2020*
Accepted: *05 February 2021*
Published: *18 March 2021*

Citation:
Ando M, Liu X, Yan Y, Yang Y, Namba S, Abe K and Kambara T (2021) Sound-Symbolic Semantics of Written Japanese Vowels in a Paper-Based Survey Study. Front. Commun. 6:617532. doi: 10.3389/fcomm.2021.617532

INTRODUCTION

In psycholinguistic studies, while the dominant phenomenon would be arbitrariness–that words are arbitrarily associated with referents (De Saussure, 1916)—the parallel or competing phenomenon, which may appear more strongly in some languages than others, is sound symbolism. This means that words, including specific vowels and consonants, are non-arbitrarily associated with specific sensorimotor and emotional features (Hamano, 1998; Imai and Kita, 2014; Sidhu and Pexman, 2019; Kawahara, 2020). Psycholinguistic studies have assumed that the occurrence of the sound-symbolic phenomena would connect to the oral shape during the production of verbal sounds (e.g., Sapir, 1929; Namba and Kambara, 2020). Thus, the sound-symbolic phenomena can occur not only in a word, but also in a verbal sound itself, like a vowel. In fact, sound-symbolic phenomena occur in both real words and pseudowords (Sidhu and Pexman, 2019). For example, specific pseudowords (e.g., *maluma* or *bouba*) are significantly associated with round figures, while other specific pseudowords

(e.g., *takete* or *kiki*) are associated with pointy figures (bouba-kiki effect, Köhler, 1947; Ramachandran and Hubbard, 2001; Westbury, 2005; Styles and Gawne, 2017; Aryani et al., 2020). In addition, French speakers associate pseudowords including round sounds or grammatically feminine endings with round shapes (Sidhu et al., 2019). The non-arbitrary relationships between pseudowords and referents also occur based on orthographically non-arbitrary relationships between pseudowords and figures. Cuskley et al. (2017) showed that both spoken and written pseudowords (e.g., *gege*) orthographically associate with meaningless figures (e.g., a round shape). Another study also found that participants quickly matched spiky frames with angular fonts (De Carolis et al., 2018). These findings suggest that while sound symbolism is not an illusion that would actually be explained by purely visual phenomena involving the shape of letters (especially since sound symbolism occurs in human communities who do not have a written system for their languages), the visual shape of written stimuli might lead to additional effects that might partially blur sound symbolic associations. Sound symbolism has been examined in behavioral experiments for word evaluation (Kambara and Umemura, 2021) and word learning (Imai et al., 2008; Kantartzis et al., 2011), and neuroscientific experiments for word comprehension (Osaka, 2011) and word learning (Asano et al., 2015; Yang et al., 2019). For instance, mismatched relationships between words (e.g., *kipi*) and referents (e.g., a round figure) increase mismatch negativity of event-related brain potentials around 400 ms after stimulus onset (N400) for 11-month-old infants compared to matched relationships between words (e.g., *moma*) and referents (e.g., a round figure; see Asano et al., 2015). Sound symbolism also facilitates word learning in first and second languages (Imai et al., 2008; Kantartzis et al., 2011; Imai and Kita, 2014;), although previous findings have also suggested that participants also arbitrarily associate novel words, including unfamiliar and meaningless words, with referents (e.g., Kambara et al., 2013; Takashima et al., 2014; Takashima et al., 2017; Liu et al., 2021). A previous study reported that a group, who implicitly learned congruently sound-symbolic relationships between pseudowords and meaningless figures, performed matching pseudowords with meaningless figures better than another group, who explicitly learned incongruently sound-symbolic relationships between pseudowords and meaningless figures (Nielsen and Rendall, 2012). The voicing of consonants in sound-symbolic words affects subjective evaluations of the words (Kambara and Umemura, 2021). Vowels influence perception of objects and faces (Sapir, 1929; Newman, 1933; Perfors, 2004) and brand names/brand personalities (Klink, 2000; Yorkston and Menon, 2004; Wu et al., 2013). Klink (2000) reported that participants evaluated brand names with front vowels (e.g., *i*) as more bitter, colder, faster, more feminine, friendlier, lighter (relative to darker), lighter (relative to heavier), milder, prettier, smaller, softer, thinner and weaker than those with back vowels (e.g., *o*). Phonemes in European languages can connect to specific emotional features (especially, affection and arousal) in behavioral (Adelman et al., 2018; Aryani and Jacobs, 2018; Aryani et al., 2018a; Myers-Schulz et al., 2013) and neural processing (Aryani et al., 2018b; Aryani et al., 2019). Although these previous studies have reported that vowels and consonants in sound-symbolic words contribute to subjective evaluations including physical and emotional features, there is still no evidence of such subjective evaluations of written vowels.

The purpose of this study was to investigate differences among subjective evaluations of Japanese vowels. Many researchers have examined Japanese sound symbolism in terms of word evaluation, word comprehension and word learning (e.g., Hamano, 1986; Hamano, 1998; Imai et al., 2008; Shinohara and Kawahara, 2010; Osaka, 2011; Imai and Kita, 2014; Asano et al., 2015; Hoshi et al., 2019; Kawahara, 2020; Motoki et al., 2020; Kambara and Umemura, 2021), while many other researchers have also examined sound symbolism in other languages, especially in European languages (e.g., Sapir, 1929; Newman, 1933; Klink, 2000; Perfors, 2004; Yorkston and Menon, 2004; Myers-Schultz et al., 2013; Wu et al., 2013; Adelman et al., 2018; Aryani and Jacobs, 2018; Aryani et al., 2018a; Ariyani et al., 2018b; Aryani et al., 2019). One of the reasons why sound symbolism researchers have focused on Japanese is the large number of sound-symbolic words in Japanese (Hamano, 1986; Hamano, 1998; Imai and Kita, 2014). One dictionary lists 4,500 sound-symbolic words in Japanese (Ono, 2007). In addition, the Japanese writing system includes five written vowels *(a*, *i*, *u*, *e,* and *o*; see Goetry et al., 2005). Thus, in this study, Japanese participants evaluated five written Japanese vowels using five-point semantic differential scales for 10 features: size, closeness, thickness, width, weight, height, depth, affection, excitement and familiarity. We made three predictions, based on previous findings (Sapir, 1929; Newman, 1933; Klink, 2000; Shinohara and Kawahara, 2010; Namba and Kambara, 2020). First, participants would evaluate that the vowel *i* is smaller than the vowel *a*. This prediction was based on previous findings that participants evaluated words with the vowel *i* as smaller than words with the vowel *a* in certain languages (Sapir, 1929; Newman, 1933; Klink, 2000; Shinohara and Kawahara, 2010). Second, the vowel *i* would be rated more preferable and more familiar than other vowels. This prediction was consistent with other studies in which words including the front vowels were evaluated more positively, for instance prettier, than those including back vowels (Klink, 2000), and the mouth shape used to produce the vowel *i* was preferable and more familiar than that used for other vowels (Namba and Kambara, 2020). Third, we also predicted that there would be other differences among physical (size, closeness, thickness, width, weight, height, and depth) or emotional evaluations (affection, excitement, and familiarity) of vowels, as well as correlations among them. Although this survey study was an exploratory approach for globally assessing the sound symbolism of different vowels in a language (Japanese), some studies have theoretically suggested differences among subjective evaluations of vowels in words (e.g., Hamano, 1986; Hamano, 1998). In addition, because psycholinguistic features of words (e.g., familiarity, imageability, affection, and excitement) correlate with each other (e.g., Kambara et al., 2020), those of vowels would also correlate with one another.

a	① small	1 2 3 4 5	big
	② far	1 2 3 4 5	close
	③ thin	1 2 3 4 5	thick
	④ narrow	1 2 3 4 5	wide
	⑤ light	1 2 3 4 5	heavy
	⑥ low	1 2 3 4 5	high
	⑦ shallow	1 2 3 4 5	deep
	⑧ dislike	1 2 3 4 5	like
	⑨ calm	1 2 3 4 5	excited
	⑩ unfamiliar	1 2 3 4 5	familiar

FIGURE 1 | An example of 10 semantic differential features for a Japanese katakana letter (ア, a). In the questionnaire, all the items were shown in Japanese.

METHODS

Participants

Six hundred and thirteen Japanese participants (482 female; M_{age} = 16.98; SD_{age} = 6.28) participated in this survey study. The participants were people that attended a lecture on introductory psychology in department of psychology at Hiroshima University. This survey study was conducted practically to introduce one example of psychological methods in the lecture. There was a majority of female participants due to enrollment in the psychology degree. Informed consent was obtained from each participant before the survey. After the lecture on introductory psychology, all participants received an A4 paper to complete the survey task voluntarily. Before the survey task, an author (TK.) explained this survey study approved by the ethical committee of the Graduate School of Education at Hiroshima University (approval code: 2019554).

Materials and Procedures

Five written Japanese vowels were used in this study. The five written Japanese vowels were associated with *a* (ア), *i* (イ), *u* (ウ), *e* (エ), and *o* (オ) sounds. The written Japanese letters were presented with Japanese katakana characters. Katakana includes a wide range of applications for the Japanese writing system, compared to other Japanese characters such as hiragana and kanji (e.g., Goetry et al., 2005). For example, in Japanese, katakana would be applied for both loanwords and non-loanwords including sound-symbolic words, while hiragana and kanji would be generally applied to non-loanwords only. Therefore, we decided to use katakana, not hiragana and kanji. The 10 semantic differential scales were based on previous studies in theoretical linguistics and psycholinguistics (Osgood et al., 1957; Hamano, 1998; Klink, 2000; Kambara et al., 2020; Namba and Kambara, 2020; Kambara and Umemura, 2021). These scales were associated with size (1: small; 5: big), closeness (1: far; 5: close), thickness (1: thin; 5: thick), width (1: narrow; 5: wide), weight (1: light; 5: heavy), height (1: low; 5: high), depth (1: shallow; 5: deep), affection (1: dislike; 5: like), excitement (1: calm; 5: excited) and familiarity (1: unfamiliar; 5: familiar). Participants evaluated each vowel on each scale, on A4 paper (**Figure 1**). The font size and style were 10.5 and Yu Mincho, respectively. The order of stimuli presentation was *a*, *i*, *u*, *e*, and *o*. The instruction at the beginning of the questionnaire was "Please select and circle a number associated with the most appropriate sensation or feeling to each katakana letter (*a*, *i*, *u*, *e*, and *o*)."

Analyses

Non-parametric one-way analysis of variance (ANOVA) tests (Friedman's tests) were applied to each scale to assess the subjective evaluations of vowels. When the ANOVA test was significant, the post hoc analyses (Wilcoxon signed rank tests) were applied using a Benjamini-Hochberg correction (Benjamini and Hochberg, 1995). With the Benjamini-Hochberg (False Discovery rate: FDR) correction, the adjusted *p*-values (*q*-values) ranged from 0.005 (the statistical threshold of the first rank) to 0.05 (the statistical threshold of the tenth rank). Listwise deletions were applied for missing values in the ANOVA tests. These analyses were performed using the SPSS software on a desktop computer running Windows. In addition, we conducted Spearman rank correlation analyses among mean scores of the 10 subjective evaluations in order to examine the relationships among the subjective evaluations, using SPSS software on a Windows-based laptop. When the Benjamini-Hochberg correction (Benjamini and Hochberg, 1995) was also applied to Spearman rank correlation analyses, the adjusted *p*-values (*q*-values) ranged from 0.001 (the statistical threshold of the first rank) to 0.05 (the statistical threshold of the forty-fifth rank). Pairwise deletions were applied for missing values in the Spearman rank correlation analyses.

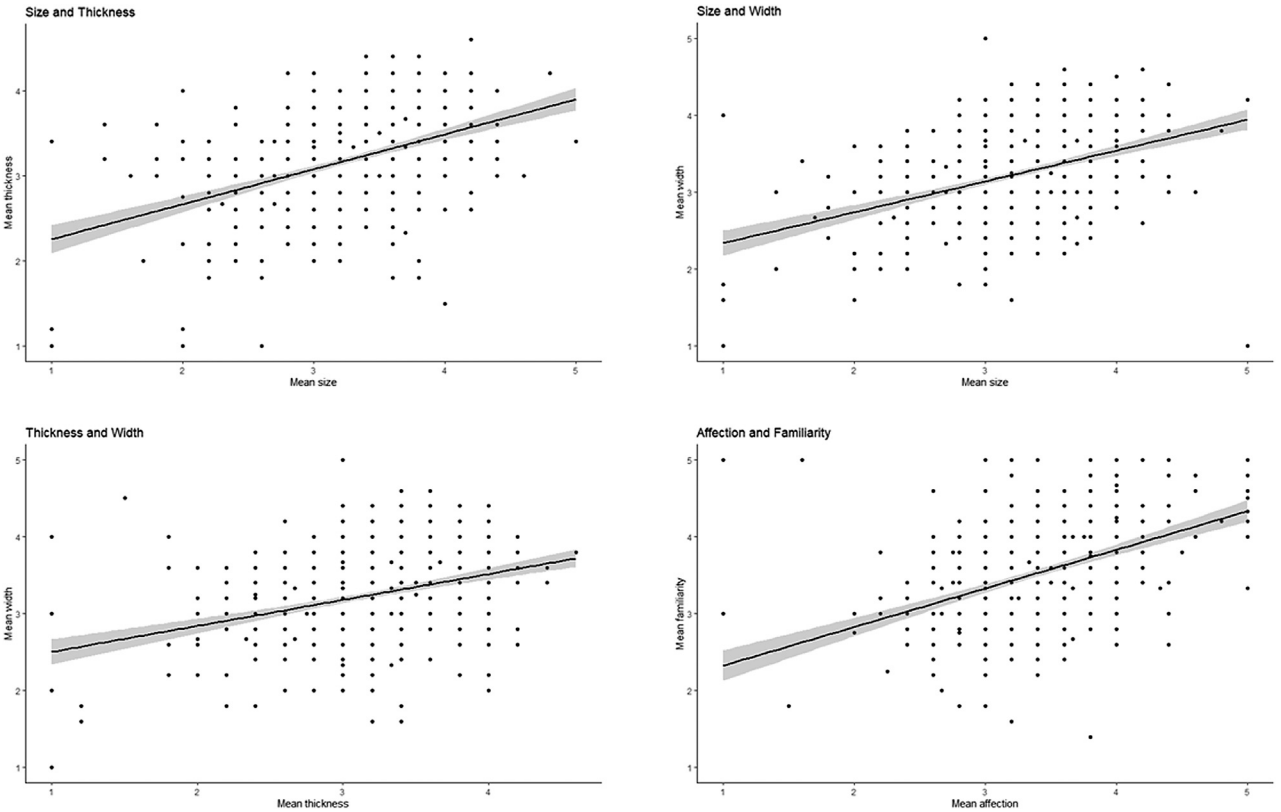


FIGURE 2 | Scatter plots for correlations between mean scores of subjective evaluations (r > 0.30). An upper left figure shows a correlation between size and thickness. An upper right figure shows a correlation between size and width. A lower left figure shows a correlation between thickness and width. A lower right figure shows a correlation between affection and familiarity.

RESULTS

Size (1: Small; 5: Big)

The Friedman's test showed that there was a significant difference among the size ratings of the five Japanese written vowels ($X_F^2(4)$ = 972.11, $p < 0.001$; **Table 1**). Medians (Inter Quartile Range: IQR) of *a*, *i*, *u*, *e*, and *o* were 4.00 (3.00–5.00), 2.00 (1.00–2.00), 3.00 (3.00–4.00), 3.00 (2.00–4.00), and 5.00 (4.00–5.00), respectively. The ratings of *o* were significantly higher than those for those of *a* ($Z = -6.69$, $p < 0.001$), *i* ($Z = -19.15$, $p < 0.001$), *u* ($Z = -11.47$, $p < 0.001$), and *e* ($Z = -13.92$, $p < 0.001$). The ratings of *a* were higher than those of *i* ($Z = -18.78$, $p < 0.001$), *u* ($Z = -5.29$, $p < 0.001$), and *e* ($Z = -10.92$, $p < 0.001$). The ratings of *u* were higher than those of *i* ($Z = -17.54$, $p < 0.001$) and *e* ($Z = -6.28$, $p < 0.001$). Finally, the ratings of *e* were higher than those of *i* ($Z = -15.27$, $p < 0.001$).

Closeness (1: Far; 5: Close)

There was a statistically significant difference among the closeness ratings of the five vowels ($X_F^2(4) = 134.52$, $p < 0.001$; **Table 1**). Medians (IQR) of *a*, *i*, *u*, *e* and *o* were 4.00 (3.00–4.00), 3.00 (2.00–4.00), 3.00 (2.00–4.00), 3.00 (2.00–4.00) and 3.00 (2.00–4.00), respectively. The ratings of *a* were higher than those of *i* ($Z = -10.29$, $p < 0.001$), *u* ($Z = -6.91$, $p < 0.001$), *e* ($Z = -9.71$, $p < 0.001$), and *o* ($Z = -5.87$, $p < 0.001$). The ratings of *u* were higher than those of *i* ($Z = -4.84$, $p < 0.001$) and *e* ($Z = -4.27$, $p < 0.001$). The ratings of *o* were higher than those of *i* ($Z = -4.18$, $p < 0.001$) and *e* ($Z = -4.16$, $p < 0.001$). No significant

TABLE 1 | Mean scores and standard deviations of subjective evaluations of five written Japanese vowels.

	M (*SD*)				
	a	***i***	***u***	***e***	***o***
Size	3.73 (1.06)	1.93 (0.88)	3.40 (1.08)	2.98 (1.15)	4.12 (1.14)
Closeness	3.55 (1.13)	2.78 (1.16)	3.12 (1.11)	2.83 (1.13)	3.13 (1.40)
Thickness	3.50 (1.06)	1.82 (0.84)	3.56 (1.06)	2.97 (1.14)	4.02 (1.03)
Width	3.68 (1.13)	2.28 (1.09)	3.39 (1.12)	3.13 (1.20)	3.71 (1.17)
Weight	3.03 (1.24)	1.85 (0.91)	3.59 (1.06)	2.98 (1.15)	4.08 (1.07)
Height	3.20 (1.13)	3.39 (1.18)	2.61 (0.99)	2.72 (1.15)	2.51 (1.29)
Depth	2.90 (1.13)	2.32 (1.02)	3.43 (0.95)	2.78 (1.12)	3.81 (1.03)
Affection	3.71 (1.00)	3.35 (0.95)	3.21 (0.98)	3.19 (0.97)	3.31 (0.98)
Excitement	2.83 (1.21)	2.79 (1.10)	2.69 (1.09)	2.81 (1.07)	2.86 (1.30)
Familiarity	4.04 (1.06)	3.31 (1.11)	3.35 (1.12)	3.27 (1.13)	3.56 (1.12)

M, mean scores; SD, standard deviation.

difference was observed between *u* and *o* ($Z = -0.26$, $p = 0.80$) or between *i* and *e* ($Z = -0.77$, $p = 0.45$).

Thickness (1: Thin; 5: Thick)

There was a statistically significant difference among the thickness ratings of the five vowels ($X_F^2(4) = 969.33$, $p < 0.001$). Medians (IQR) of *a*, *i*, *u*, *e*, and *o* were 4.00 (3.00–4.00), 2.00 (1.00–2.00), 4.00 (3.00–4.00), 3.00 (2.00–4.00), and 4.00 (3.00–5.00), respectively. The ratings of *o* were higher than those of *a* ($Z = -8.99$, $p < 0.001$), *i* ($Z = -19.39$, $p < 0.001$), *u* ($Z = -8.04$, $p < 0.001$) and *e* ($Z = -13.90$, $p < 0.001$). The ratings of *u* were higher than those of *i* ($Z = -18.56$, $p < 0.001$) and *e* ($Z = -8.98$, $p < 0.001$). In addition, the ratings of *a* were higher than those of *i* ($Z = -18.50$, $p < 0.001$) and *e* ($Z = -8.09$, $p < 0.001$). Finally, the ratings of *e* were higher than those of *i* ($Z = -15.43$, $p < 0.001$). There was no significant difference between *a* and *u* ($Z = -0.97$, $p = 0.33$).

Width (1: Narrow; 5: Wide)

There was a statistically significant difference among the width ratings of the five vowels ($X_F^2(4) = 441.41$, $p < 0.001$; **Table 1**). Medians (IQR) of *a*, *i*, *u*, *e*, and *o* were 4.00 (3.00–5.00), 2.00 (1.00–3.00), 3.00 (3.00–4.00), 3.00 (2.00–4.00), and 4.00 (3.00–5.00), respectively. The ratings of *o* were higher than those of *i* ($Z = -15.06$, $p < 0.001$), *u* ($Z = -4.95$, $p < 0.001$) and *e* ($Z = -7.96$, $p < 0.001$). The ratings of *a* were also higher than those of *i* ($Z = -15.44$, $p < 0.001$), *u* ($Z = -4.09$, $p < 0.001$) and *e* ($Z = -7.93$, $p < 0.001$). The ratings of *u* were higher than those of *i* ($Z = -14.39$, $p < 0.001$) and *e* ($Z = -4.03$, $p < 0.001$). In addition, the ratings of *e* were higher than those of *i* ($Z = -11.43$, $p < 0.001$). There was no significant difference between the ratings of *a* and *o* ($Z = -0.45$, $p = 0.65$).

Weight (1: Light; 5: Heavy)

There was a statistically significant difference among the weight ratings of the five vowels ($X_F^2(4) = 880.32$, $p < 0.001$; **Table 1**). Medians (IQR) of *a*, *i*, *u*, *e*, and *o* were 3.00 (2.00–4.00), 2.00 (1.00–2.00), 4.00 (3.00–4.00), 3.00 (2.00–4.00), and 4.00 (3.00–5.00), respectively. The rating of *o* was higher than those of *a* ($Z = -13.36$, $p < 0.001$), *i* ($Z = -19.03$, $p < 0.001$), *u* ($Z = -8.43$, $p < 0.001$), and *e* ($Z = -13.90$, $p < 0.001$). The ratings of *u* were higher than those of *a* ($Z = -8.24$, $p < 0.001$), *i* ($Z = -18.38$, $p < 0.001$), and *e* ($Z = -8.71$, $p < 0.001$). In addition, the ratings of *a* were higher than those of *i* ($Z = -14.77$, $p < 0.001$). Finally, the ratings of *e* were higher than those of *i* ($Z = -15.23$, $p < 0.001$). There was no significant difference between *a* and *e* ($Z = -0.64$, $p = 0.53$).

Height (1: Low; 5: High)

There was a statistically significant difference among height ratings of five vowels ($X_F^2(4) = 224.59$, $p < 0.001$; **Table 1**). Medians (IQR) of *a*, *i*, *u*, *e*, and *o* were 3.00 (2.00–4.00), 4.00 (3.00–4.00), 3.00 (2.00–3.00), 3.00 (2.00–4.00), and 2.00 (1.00–3.00), respectively. The ratings of *i* were higher than those of the other vowels, including *a* ($Z = -3.15$, $p < 0.005$), *u* ($Z = -11.00$, $p < 0.001$), *e* ($Z = -9.61$, $p < 0.001$) and *o* ($Z = -10.42$, $p < 0.001$). The ratings of *a* were higher than those of *u* ($Z = -9.36$, $p < 0.001$), *e* ($Z = -6.87$, $p < 0.001$) and *o* ($Z = -9.62$, $p < 0.001$). The ratings of *e* were higher than those of *o* ($Z = -2.68$, $p = 0.007$). The other comparisons (*u* vs. *e*, $Z = -1.56$, $p = 0.12$; *u* vs. *o*, $Z = -1.59$, $p = 0.11$) were not satisfied for the statistical threshold corrected with the Benjamini-Hochberg (FDR) method.

Depth (1: Shallow; 5: Deep)

There was a statistically significant difference among depth ratings of five vowels ($X_F^2(4) = 546.92$, $p < 0.001$; **Table 1**). Medians (IQR) of *a*, *i*, *u*, *e*, and *o* were 3.00 (2.00–4.00), 2.00 (2.00–3.00), 3.00 (3.00–4.00), 3.00 (2.00–4.00), and 4.00 (3.00–5.00), respectively. The ratings of *o* were higher (deeper) than those of *a* ($Z = -12.44$, $p < 0.001$), *i* ($Z = -16.01$, $p < 0.001$), *u* ($Z = -6.74$, $p < 0.001$), and *e* ($Z = -13.37$, p 0.001). The ratings of *u* were also higher than those of *a* ($Z = -8.77$, $p < 0.001$), *i* ($Z = -15.20$, $p < 0.001$), and *e* ($Z = -10.04$, $p < 0.001$). The ratings of *a* were also higher than those of *i* ($Z = -9.15$, $p < 0.001$). The ratings of *e* were also higher than those of *i* ($Z = -7.21$, $p < 0.001$). In addition, a comparison between the ratings of *a* and *e* ($Z = -1.87$, $p = 0.06$) was not satisfied for the statistical threshold corrected with the Benjamini-Hochberg (FDR) method.

Affection (1: Dislike; 5: Like)

There was a statistically significant difference among affection ratings of five vowels ($X_F^2(4) = 144.45$, $p < 0.001$; **Table 1**). Medians (IQR) of *a*, *i*, *u*, *e*, and *o* were 4.00 (3.00–5.00), 3.00 (3.00–4.00), 3.00 (3.00–4.00), 3.00 (3.00–4.00), and 3.00 (3.00–4.00), respectively. The ratings of *a* were higher (more preferable) than those of the other vowels: *i* ($Z = -6.63$, $p < 0.001$), *u* ($Z = -8.82$, $p < 0.001$), *e* ($Z = -8.79$, $p < 0.001$), and *o* ($Z = -7.31$, $p < 0.001$). The ratings of *i* were higher than those of *u* ($Z = -2.64$, $p = 0.008$) and *e* ($Z = -3.10$, $p < 0.005$). In addition, the ratings of *o* were also higher than those of *e* ($Z = -2.16$, $p = 0.03$). The other comparisons (*i* vs. *o*, $Z = -0.50$, $p = 0.62$; *u* vs. *e*, $Z = -0.24$, $p = 0.81$; *u* vs. *o*, $Z = -1.77$, $p = 0.08$) were not satisfied for the statistical threshold corrected with the Benjamini-Hochberg (FDR) method.

Excitement (1: Calm; 5: Excited)

There was not a statistically significant difference among the excitement ratings of the five vowels ($X_F^2(4) = 5.58$, $p = 0.23$; **Table 1**). Medians (IQR) of *a*, *i*, *u*, *e*, and *o* were 3.00 (2.00–4.00), 3.00 (2.00–4.00), 3.00 (2.00–3.00), 3.00 (2.00–4.00), and 3.00 (2.00–4.00), respectively.

Familiarity (1: Unfamiliar; 5: Familiar)

There was a statistically significance among familiarity ratings of five vowels ($X_F^2(4) = 222.29$, $p < 0.001$; **Table 1**). Medians (IQR) of *a*, *i*, *u*, *e*, and *o* were 4.00 (3.00–5.00), 3.00 (3.00–4.00), 3.00 (3.00–4.00), 3.00 (2.00–4.00), and 4.00 (3.00–4.00), respectively. The ratings of *a* were higher (more familiar) than those of *i* ($Z = -10.79$, $p < 0.001$), *u* ($Z = -10.30$, $p < 0.001$), *e* ($Z = -11.48$, $p < 0.001$) and *o* ($Z = -7.82$, $p < 0.001$). In addition, the rating of *o* was also higher (more familiar) than those of *i* ($Z = -4.04$, $p < 0.001$), *u* ($Z = -3.69$, $p < 0.001$) and *e* ($Z = -5.02$, $p < 0.001$). The other comparisons (*i* vs. *u*, $Z = -0.82$, $p = 0.41$; *i* vs. *e*, $Z = -0.89$, $p = 0.37$; *u* vs. *e*, $Z = -1.71$, $p = 0.09$) were not satisfied for the statistical threshold corrected with the Benjamini-Hochberg (FDR) method.

TABLE 2 | Results of Spearman rank correlation analyses among mean scores of each subjective evaluation.

	Size	Closeness	Thickness	Width	Weight	Height	Depth	Affection	Excitement	Familiarity
Size										
Closeness	0.18[a]									
Thickness	0.40[a]	0.12[a]								
Width	0.41[a]	0.14[a]	0.34[a]							
Weight	0.21[a]	0.06	0.29[a]	0.20[a]						
Height	−0.04	0.06	−0.11[a]	−0.04	−0.12[a]					
Depth	0.24[a]	0.03	0.17[a]	0.13[a]	0.26[a]	−0.01				
Affection	0.09	0.15[a]	0.09[a]	0.08	−0.10[a]	0.08	−0.03			
Excitement	−0.04	0.01	0.00	−0.05	0.06	0.00	0.07	−0.13[a]		
Familiarity	0.04	0.14[a]	0.05	0.11[a]	−0.03	0.09	0.06	0.43[a]	-0.11[a]	

[a]*A significant correlation corrected by Benjamini-Hochberg procedure (Benjamini and Hochberg, 1995).*

Correlations Among the Mean Scores of Subjective Evaluations

There were significant correlations among the mean scores of subjective evaluations of vowels. All significant correlations are shown in **Table 2**. The effect sizes (*rs*) of four significant positive correlations between subjective evaluations (size and thickness; size and width; thickness and width; affection and familiarity) ranged between 0.30 (medium effect size) and 0.50 (large effect size; Cohen, 1988). In addition, the effect sizes (rs) of all the other significant positive correlations between subjective evaluations (size and closeness; size and weight; size and depth; closeness and thickness; closeness and width; closeness and affection; closeness and familiarity; thickness and weight; thickness and depth; width and weight; width and depth; width and familiarity; weight and depth) ranged from 0.10 (small effect size) to 0.30 (medium effect size; Cohen, 1988), except for thickness and affection, while the effect sizes (*rs*) of the significant negative correlations between subjective evaluations (thickness and height; weight and height; weight and affection; affection and excitement; excitement and familiarity) also ranged from 0.10 (small effect size) to 0.30 (medium effect size; Cohen, 1988). Although the effect size (*r*) of the significant positive correlation between thickness and affection was lower than 0.10 (small effect size; Cohen, 1988), the positive correlation satisfied the statistical threshold of the Benjamini-Hochberg (FDR) correction.

DISCUSSION

The aim of the current study was to identify whether each written Japanese vowel was individually associated with specific subjective evaluations on 10 semantic differential scales (size, closeness, thickness, width, weight, height, depth, affection, excitement, and familiarity). Japanese native speakers rated the 10 features on five-point semantic differential scales for five presented written Japanese vowels (*a*, *i*, *u*, *e*, and *o*). The results showed that the size, closeness, thickness and width of *a*, *u*, and *o* were significantly higher than those of *i* and *e*, whereas the affection and familiarity of *a* was significantly higher than for the other vowels. In addition, we found correlations among the mean scores of subjective evaluations. These findings suggest that each written Japanese vowel could individually contribute to specific subjective evaluations. Furthermore, current findings contribute to developing a theory of orthographically non-arbitrary associations between new spoken or written letters and referents (e.g., Cuskley et al., 2017).

Subjective Evaluation of Physical Features

We found that the size, closeness, thickness and width of *a*, *u*, and *o* were significantly higher than those of *i* and *e*, as noted. In addition, there were significant correlations among the mean scores of the subjective evaluations. These findings were consistent with previous findings of sound symbolism. Based on previous studies of sound symbolism, vowels with high second formant (F2) are *i* and *e*, while vowels with low F2 are *a*, *u*, and *o* (Ohala, 1994; Berlin, 2006; Nishi et al., 2008). Vowels with high F2 are called front vowels (*i* and *e*), whereas low-F2 vowels are called back vowels (*a*, *u*, and *o*; see Berlin, 2006). Previous studies have reported that words including front vowels were 'smaller' than words including back vowels in certain languages (Sapir, 1929; Newman, 1933; Ohala, 1994; Klink, 2000; Berlin, 2006; Shinohara and Kawahara, 2010) and that size evaluations of vowels in words are associated with the size of the oral cavity or mouth shape when pronouncing them (Sapir, 1929; Ohala, 1984; Ohala, 1994; Shinohara and Kawahara, 2010; Namba and Kambara, 2020). Ohala (1994) called such associations between high acoustic frequency and smallness, and associations between low acoustic frequency and largeness, the "frequency code." The current study suggests that physical evaluations (size, closeness, thickness and width) of written individual vowels could also be associated with the size of the oral cavity or mouth shape used to produce the vowels.

Regarding the relationship between stimulus effects and subjective evaluations to written vowels, the results of this study would also be affected by the orthographical features of the stimuli. The verbal stimuli of this study were written in katakana vowels in Japanese (i.e., *a*: ア; *i*: イ; *u*: ウ; *e*: エ; and *o*: オ), which might be in particular more angular than other Japanese characters (e.g., in hiragana, which is one of the Japanese character scripts, *a*: あ; *i*: い; *u*: う; *e*: え; and *o*: お). For instance, regarding subjective evaluations of physical features, participants might feel that the orthographical shape of *i* (イ) is thinner than the others. In fact, they judged the thickness ratings

of *i* as lower (thinner) than the others. A study supports orthographically non-arbitrary associations between verbal stimuli and referents. Cuskley et al., (2017) prepared written and spoken pseudowords with curved graphemes (e.g., *dede*) and written and spoken pseudowords with angular graphemes (e.g., *zeze*) as verbal stimuli, while they also prepared round figures and spiky figures as figure stimuli (Cuskley et al., 2017). Participants judged how well a spoken or written pseudoword matches with a presented figure in their experiments (Cuskley et al., 2017). They found that both spoken and written pseudowords (e.g., *dede*) orthographically and non-arbitrarily associate (match) with meaningless figures (e.g., round shapes; see Cuskley et al., 2017). Grapheme-color synesthesia would also support orthographical associations between letters and referents (e.g., Rouw and Scholte, 2007; Asano and Yokosawa, 2013). Since the orthographical shapes of written stimuli might include additional effects that might partially blur sound symbolic associations, future studies also need to use spoken vowels, and examine orthographically non-arbitrary associations between spoken vowels and referents (subjective evaluations of physical and emotional features).

Our findings were also associate with the stimulus presentation order. In this paper-based survey study, the order of stimulus presentation was fixed. The presentation order of the vowels was *a*, *i*, *u*, *e*, and *o*, respectively. The fixed presentation order of stimuli might affect the results of this study. At least, effects of stimulus presentation order occur in spoken stimuli (Francis and Ciocca, 2003). Although randomizing the presentation of vowels does not guarantee an absence of such meta-strategies, it controls for the effects of the presentation order. If the order of presentation were randomized, order effects could be controlled for.

In addition, although previous findings have shown that both vowels or consonants in sound symbolic words and pseudowords associate with referential features (e.g., Klink, 2000; Cuskley et al., 2017; Kambara and Umemura, 2021), current findings focus on how an isolated written vowel associates with subjective evaluations of physical and emotional features. From here, vowels occurring in words or pseudowords could be further investigated. Also, since consonants in sound symbolic words and pseudowords have also been shown to associate with specific evaluations of physical and emotional features (e.g., Klink, 2000; Cuskley et al., 2017; Kambara and Umemura, 2021), interactions between vowels and consonants may occur in words and be further studied. Japanese might be a specific case here, since most katakana and hiragana characters associate a consonant and a vowel (Goetry et al., 2005). The comparison between katakana and hiragana could also be meaningful, and approached in future studies with the methods of this study.

Finally, as we mentioned in the Methods, all participants received an A4 paper to perform the survey task voluntarily after a lecture of introductory psychology. The participants might have directly compared between their subjective evaluations of vowels with this survey method. If so, their evaluations could reflect such meta-strategies of comparison. Thus, the explicit nature of judgments in this study might have affected the findings (e.g., Nielsen and Rendall, 2012). On the other hand, the participants might feel mental or physical fatigues before the survey study. Additionally, their mental or physical fatigues might affect subjective evaluations of vowels. In addition, since the majority of participants were female, the sample composition might have affected the findings of the current study.

Subjective Evaluation of Emotional Features

Affection and familiarity for the vowel *a* were higher than those of the other vowels (*i*, *u*, *e*, and *o*). In addition, the mean scores of affection positively correlated with those of familiarity. These results are associated with the order of vowels in the Japanese writing system, where *a* comes first of all vowels and all letters (Goetry et al., 2005), although *a* is also the first letter (vowel) in the Latin alphabet. Stockman et al. (1981) found that listeners could detect low front vowels including (e.g., /a/), back consonants (e.g., /h/), and schwa /ə/, which were about 70 percent of all the transcribed data on early baby vocalization from 7 to 21 months old, although vowels which infants produce in babbling would change in each period of infant development (Smith and Oller, 1981). Phonological segments *a* and *i* in languages have high frequency in a repository of cross-linguistically phonological inventory data (Moran and McCloy, 2019; see https://phoible.org/). The letter *a* might be cross-linguistically familiar and detectable for people. In addition, the mean scores of affection and familiarity correlated positively with those of closeness. Since the subjective evaluations of closeness are associated with other physical evaluations, the initial effect of the order of the vowels in the Japanese writing system might also be associated with the physical evaluation of vowel sounds. However, our findings were inconsistent with previous findings in which words with front vowels (e.g., *i*) were evaluated as associated with more positive emotions (e.g., prettier) than were those with back vowels (e.g., *o*; see Klink, 2000), and also in which the mouth shape used to pronounce a front vowel (e.g., *i*) was considered connected with being more positive, calm, and familiar about pronouncing a front vowel than that with a back vowel (e.g., *o*; Namba and Kambara, 2020). Although Ohala (1984) also claimed facial expressions sound-symbolically affect relationships between vowels and emotional features, the current findings did not support these previous studies. The first possible reason for this inconsistency between this and previous findings might be related to differences between languages. Indeed, some studies do not support universal sound symbolism (Maltzman et al., 1956; Brackbill and Little, 1957; Atzet and Gerard, 1965). The phonological information of words associates with word classes (e.g., nouns and verbs), but the phonological information is differently distributed in different languages (Monaghan et al., 2007). Thus, the current findings of written Japanese vowels might only show non-universal sound symbolic associations. The second possible reason for inconsistency between current and previous studies might be whether the presented stimuli were words (Klink, 2000), mouth shapes to pronounce vowels (Namba and Kambara, 2020), or vowels alone in this study. For example, in cases where the presented stimuli are full words, they would include both consonants and vowels, and the consonants in the

words would involve their own sound-symbolic effects (e.g., Klink, 2000; Kambara and Umemura, 2021), which might in turn affect subjective evaluations of vowels in the words.

CONCLUSION

We examined whether each written Japanese vowel was associated with specific subjective features by using 10 semantic differential scales, respectively assessing subjective size, closeness, thickness, width, weight, height, depth, affection, excitement and familiarity. Japanese native speakers rated each written Japanese vowel (*a*, *i*, *u*, *e*, and *o*) on each five-point semantic differential scale. We found that the size, closeness, thickness and width of *a*, *u*, and *o* were significantly higher than those of *i* and *e*, whereas the affection and familiarity of *a* were higher than those of the other vowels (*i*, *u*, *e*, and *o*). We also found correlations among the mean scores of subjective evaluations. Taken together, these findings suggest that each written Japanese vowel individually contributes to specific subjective evaluations.

DATA AVAILABILITY STATEMENT

The datasets presented in this article are not readily available because the data include personal information. Requests to access the datasets should be directed to Toshimune Kambara, tkambara@hiroshima-u.ac.jp.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethical Committee of the Graduate School of Education at Hiroshima University. Written informed consent from the participants' legal guardian/next of kin was not required in accordance with the national legislation and the institutional requirements.

AUTHOR CONTRIBUTIONS

Conceptualization, TK; methodology, MA, XL, YY, YY, SN, and TK; software, MA and TK; validation, MA; formal analyses, MA and TK; investigation, SN, KA and TK; resources, TK; data curation, MA, XL, YY, and YY; writing—original draft preparation, MA. and TK.; writing—review and editing, SN, KA, and TK; visualization, MA and TK; supervision, TK; project administration, TK; funding acquisition, TK All authors have read and agreed to the published version of the manuscript.

FUNDING

The corresponding author (TK) was supported by Hiroshima University Grant-in-Aid for Scientific Research, KAKENHI Grant-in-Aid for Research Activity Start-up, KAKENHI Grant-in-Aid for Early-Career Scientists and KAKENHI Grant-in-Aid for Scientific Research (C). In addition, this research was conducted as part of the School of Education Joint Research Project 2020 at Hiroshima University, and received research support from the School of Education.

ACKNOWLEDGMENTS

We would like to thank editors and reviewers who provided important suggestions to improve this article.

REFERENCES

Adelman, J. S., Estes, Z., and Cossu, M. (2018). Emotional sound symbolism: languages rapidly signal valence *via* phonemes. *Cognition* 175, 122–130. doi:10.1016/j.cognition.2018.02.007

Aryani, A., Conrad, M., Schmidtke, D., and Jacobs, A. (2018a). Why "piss" is ruder than "pee"? The role of sound in affective meaning making. *PLoS One* 13 (6), e0198430. doi:10.1371/journal.pone.0198430

Aryani, A., Hsu, C.-T., and Jacobs, A. M. (2018b). The sound of words evokes affective brain responses. *Brain Sci.* 8 (6), 94. doi:10.3390/brainsci8060094

Aryani, A., Hsu, C. T., and Jacobs, A. M. (2019). Affective iconic words benefit from additional sound-meaning integration in the left amygdala. *Hum. Brain Mapp.* 40 (18), 5289–5300. doi:10.1002/hbm.24772

Aryani, A., Isbilen, E. S., and Christiansen, M. H. (2020). Affective arousal links sound to meaning. *Psychol. Sci.* 31 (8), 978–986. doi:10.1177/0956797620927967

Aryani, A., and Jacobs, A. M. (2018). Affective congruence between sound and meaning of words facilitates semantic decision. *Behav. Sci.* 8 (6), 56. doi:10.3390/bs8060056

Asano, M., Imai, M., Kita, S., Kitajo, K., Okada, H., and Thierry, G. (2015). Sound symbolism scaffolds language development in preverbal infants. *Cortex* 63, 196–205. doi:10.1016/j.cortex.2014.08.025

Asano, M., and Yokosawa, K. (2013). Grapheme learning and grapheme-color synesthesia: toward a comprehensive model of grapheme-color association. *Front. Hum. Neurosci.* 7, 757. doi:10.3389/fnhum.2013.00757

Atzet, J., and Gerard, H. B. (1965). A study of phonetic symbolism among native Navajo speakers. *J. Personal. Soc. Psychol.* 1 (5), 524–528. doi:10.1037/h0021882

Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B (Methodological)* 57 (1), 289–300. doi:10.1111/j.2517-6161.1995.tb02031.x

Berlin, B. (2006). The first congress of ethnozoological nomenclature. *J. R. Anthropol. Inst.* 12 (1), S23–S44. doi:10.1111/j.1467-9655.2006.00271.x

Brackbill, Y., and Little, K. B. (1957). Factors determining the guessing of meanings of foreign words. *J. Abnorm. Psychol.* 54 (3), 312–318. doi:10.1037/h0042411

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Lawrence Erlbaum.

Cuskley, C., Simner, J., and Kirby, S. (2017). Phonological and orthographic influences in the bouba-kiki effect. *Psychol. Res.* 81 (1), 119–130. doi:10.1007/s00426-015-0709-2

De Carolis, L., Marsico, E., Arnaud, V., and Coupé, C. (2018). Assessing sound symbolism: investigating phonetic forms, visual shapes and letter fonts in an implicit bouba-kiki experimental paradigm. *PLoS One* 13 (12), e0208874. doi:10.1371/journal.pone.0208874

De Saussure, F. (1916). *Course in general linguistics*. New York, NY: Columbia University Press.

Francis, A. L., and Ciocca, V. (2003). Stimulus presentation order and the perception of lexical tones in cantonese. *J. Acoust. Soc. Am.* 114 (3), 1611–1621. doi:10.1121/1.1603231

Goetry, V., Urbain, S., Morais, J., and Kolinsky, R. (2005). Paths to phonemic awareness in Japanese: evidence from a training study. *Appl. Psycholinguistics* 26, 285–309. doi:10.1017/S0142716405050186

Hamano, S. (1986). The sound-symbolic system of japanese. Ph. D. Dissertation. Gainesville (FL): University of Florida.

Hamano, S. (1998). The sound-symbolic system of Japanese. *Stud. Jap. Ling.* 10, 1–262.

Hoshi, H., Kwon, N., Akita, K., and Auracher, J. (2019). Semantic associations dominate over perceptual associations in vowel-size iconicity. *Iperception* 10, 2041669519861981. doi:10.1177/2041669519861981

Imai, M., Kita, S., Nagumo, M., and Okada, H. (2008). Sound symbolism facilitates early verb learning. *Cognition* 109, 54–65. doi:10.1016/j.cognition.2008.07.015

Imai, M., and Kita, S. (2014). The sound symbolism bootstrapping hypothesis for language acquisition and language evolution. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 369, 20130298. doi:10.1098/rstb.2013.0298

Kambara, T., Tsukiura, T., Shigemune, Y., Kanno, A., Nouchi, R., Yomogida, Y., et al. (2013). Learning-dependent changes of associations between unfamiliar words and perceptual features: a 15-day longitudinal study. *Lang. Sci.* 35, 80–86. doi:10.1016/j.langsci.2012.05.001

Kambara, T., and Umemura, T. (2021). The relationships between initial consonants in Japanese sound symbolic words and familiarity, multi-sensory imageability, emotional valence, and arousal. *J. Psycholinguist Res.* doi:10.1007/s10936-020-09749-w

Kambara, T., Umemura, T., Ackert, M., and Yang, Y. (2020). The relationship between psycholinguistic features of religious words and core dimensions of religiosity: a survey study with Japanese participants. *Religions* 11 (12), 673. doi:10.3390/rel11120673

Kantartzis, K., Imai, M., and Kita, S. (2011). Japanese sound-symbolism facilitates word learning in English-speaking children. *Cognit. Sci.* 35, 575–586. doi:10.1111/j.1551-6709.2010.01169.x

Kawahara, S. (2020). Sound symbolism and theoretical phonology. *Lang. Ling. Compass* 14, e12372. doi:10.1111/lnc3.12372

Klink, R. R. (2000). Creating brand names with meaning: the use of sound symbolism. *Mark. Lett.* 11 (1), 5–20. doi:10.1023/A:1008184423824

Köhler, W. (1947). *Gestalt psychology*. New York, NY: Liveright.

Liu, X., Horinouchi, H., Yang, Y., Yan, Y., Ando, M., Obinna, U. J., et al. (2021). Pictorial referents facilitate recognition and retrieval speeds of associations between novel words in a second language (L2) and referents. *Front. Commun.* doi:10.3389/fcomm.2021.605009

Maltzman, I., Morrisett, L. J., Jr., and Brooks, L. O. (1956). An investigation of phonetic symbolism. *J. Abnormal Psychol.* 53 (2), 249–251. doi:10.1037/h0048406

Monaghan, P., Christiansen, M. H., and Chater, N. (2007). The phonological-distributional coherence hypothesis: cross-linguistic evidence in language acquisition. *Cogn. Psychol.* 55 (4), 259–305. doi:10.1016/j.cogpsych.2006.12.001

S. Moran and D. McCloy (Editors) (2019). *Phoible 2.0*. Jena, Germany: Max Planck Institute for the Science of Human History. Available at: http://phoible.org (Accessed February 22, 2021).

Motoki, K., Saito, T., Park, J., Velasco, C., Spence, C., and Sugiura, M. (2020). Tasting names: systematic investigations of taste-speech sounds associations. *Food Qual. Preference* 80, 103801. doi:10.1016/j.foodqual.2019.103801

Myers-Schulz, B., Pujara, M., Wolf, R. C., and Koenigs, M. (2013). Inherent emotional quality of human speech sounds. *Cogn. Emot.* 27 (6), 1105–1113. doi:10.1080/02699931.2012.754739

Namba, S., and Kambara, T. (2020). Semantics based on the physical characteristics of facial expressions used to produce Japanese vowels. *Behav. Sci.* 10 (10), 157. doi:10.3390/bs10100157

Newman, S. S. (1933). Further experiments in phonetic symbolism. *Am. J. Psychol.* 45 (1), 53–75. doi:10.2307/1414186

Nielsen, A., and Rendall, D. (2012). The source and magnitude of sound-symbolic biases in processing artificial word material and their implications for language learning and transmission. *Lang. Cognit.* 4 (2), 115–125. doi:10.1515/langcog-2012-0007

Nishi, K., Strange, W., Akahane-Yamada, R., Kubo, R., and Trent-Brown, S. A. (2008). Acoustic and perceptual similarity of Japanese and American English vowels. *J. Acoust. Soc. Am.* 124, 576–588. doi:10.1121/1.2931949

Ohala, J. J. (1984). An ethological perspective on common cross-language utilization of F0 of voice. *Phonetica* 41, 1–16. doi:10.1159/000261706

Ohala, J. J. (1994). "The frequency code underlies the sound symbolic use of voice pitch," in *Sound symbolism*. Editors L. Hinton, J. Nichols, and J. J. Ohala (Cambridge, United Kingdom: Cambridge University Press), 325–347.

M. Ono (Editors) (2007). *Nihongo onomatope jiten (Japanese onomatopoeia dictionary)*. Tokyo, Japan: Shogakukan.

Osaka, N. (2011). Ideomotor response and the neural representation of implied crying in the human brain: an fMRI study using onomatopoeia. *Japanese Psychol. Res.* 53 (4), 372–378. doi:10.1111/j.1468-5884.2011.00489.x

Osgood, C. E., Suci, G., and Tannenbaum, P. (1957). *The measurement of meaning*. Urbana, IL: University of Illinois Press.

Perfors, A. (2004). "What's in a name?: the effect of sound symbolism on perception of facial attractiveness," in Proceedings of the twenty-sixth annual conference of the cognitive science society. Editors K. Forbus, D Gentner, and T Regier, Chicago, IL, August 4–7, 2004 (Chicago, IL: Cognitive Science Society), 1617.

Ramachandran, V. S., and Hubbard, E. M. (2001). Synaesthesia: a window into perception, thought, and language. *J. Conscious. Stud.* 8 (12), 3–34.

Rouw, R., and Scholte, H. S. (2007). Increased structural connectivity in grapheme-color synesthesia. *Nat. Neurosci.* 10, 792–797. doi:10.1038/nn1906

Sapir, E. (1929). A study in phonetic symbolism. *J. Exp. Psychol.* 12, 225–239. doi:10.1037/h0070931

Shinohara, K., and Kawahara, S. (2010). "A cross-linguistic study of sound symbolism," in Proceedings of the 36th annual meeting of the berkeley linguistics society, Berkeley, CA, February 6–7, 2010. Berkeley, CA: University of California Press, 396–410.

Sidhu, D. M., and Pexman, P. M. (2019). The sound symbolism of names. *Curr. Dir. Psychol. Sci.* 28 (4), 398–402. doi:10.1177/0963721419850134

Sidhu, D. M., Pexman, P. M., and Saint-Aubin, J. (2019). Is un stylo sharper than une épée? Investigating the interaction of sound symbolism and grammatical gender in English and French speakers. *PLoS One* 14 (12), e0225623. doi:10.1371/journal.pone.0225623

Smith, B. L., and Oller, D. K. (1981). A comparative study of pre-meaningful vocalizations produced by normally developing and down's syndrome infants. *J. Speech Hear. Disord.* 46 (1), 46–51. doi:10.1044/jshd.4601.46

Stockman, I. J., Woods, D. R., and Tishman, A. (1981). Listener agreement on phonetic segments in Early infant vocalizations. *J. Psycholinguist Res.* 10 (6), 593–617. doi:10.1007/BF01067296

Styles, S. J., and Gawne, L. (2017). When does Maluma/Takete fail? two key failures and a meta-analysis suggest that phonology and phonotactics matter. *Iperception* 8 (4), 2041669517724807. doi:10.1177/2041669517724807

Takashima, A., Bakker, I., van Hell, J. G., Janzen, G., and McQueen, J. M. (2014). Richness of information about novel words influences how episodic and semantic memory networks interact during lexicalization. *NeuroImage* 84, 265–278. doi:10.1016/j.neuroimage.2013.08.023

Takashima, A., Bakker, I., van Hell, J. G., Janzen, G., and McQueen, J. M. (2017). Interaction between episodic and semantic memory networks in the acquisition and consolidation of novel spoken words. *Brain Lang.* 167, 44–60. doi:10.1016/j.bandl.2016.05.009

Westbury, C. (2005). Implicit sound symbolism in lexical access: evidence from an interference task. *Brain Lang.* 93, 10–19. doi:10.1016/j.bandl.2004.07.006

Wu, L., Klink, R. R., and Guo, J. (2013). Creating gender brand personality with brand names: the effects of phonetic symbolism. *J. Marketing Theory Pract.* 21 (3), 319–330. doi:10.2753/MTP1069-6679210306

Yang, J., Asano, M., Kanazawa, S., Yamaguchi, M. K., and Imai, M. (2019). Sound symbolism processing is lateralized to the right temporal region in the prelinguistic infant brain. *Sci. Rep.* 9, 13435. doi:10.1038/s41598-019-49917-0

Yorkston, E., and Menon, G. (2004). A sound idea: phonetic effects of brand names on consumer judgments. *J. Consum. Res.* 31, 43–51. doi:10.1086/383422

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Ando, Liu, Yan, Yang, Namba, Abe and Kambara. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Psychology

ORIGINAL RESEARCH
published: 31 May 2021
doi: 10.3389/fpsyg.2021.685351

The Impacts of Perceived Risk and Negative Emotions on the Service Recovery Effect for Online Travel Agencies: The Moderating Role of Corporate Reputation

*Jiahua Wei**

Business School, Guilin University of Technology, Guilin, China

This study explores the impact mechanism of perceived risk and negative emotions on the service recovery effect of an online travel agency (OTA) through a scenario experiment. The results show that: perceived risk has positive and negative impacts on negative emotions and service recovery satisfaction, negative emotions have a negative impact on service recovery satisfaction, and corporate reputation plays a positive moderating role in the relationship between perceived risk and service recovery satisfaction. This study is helpful to better explain the impact mechanism of the service recovery effect of OTAs, and to provide a theoretical reference for improving the service recovery effect of OTAs.

Keywords: OTA, service recovery effect, perceived risk, negative emotions, corporate reputation

OPEN ACCESS

Edited by:
Francesca D'Errico,
University of Bari Aldo Moro, Italy

Reviewed by:
Christian Cecconi,
Sapienza University of Rome, Italy
Ernestina Lamponi,
Università Telematica Internazionale Uninettuno, Italy

***Correspondence:**
Jiahua Wei
jiahua6688@163.com

Specialty section:
This article was submitted to Emotion Science, a section of the journal Frontiers in Psychology

Received: *25 March 2021*
Accepted: *03 May 2021*
Published: *31 May 2021*

Citation:
Wei J (2021) The Impacts of Perceived Risk and Negative Emotions on the Service Recovery Effect for Online Travel Agencies: The Moderating Role of Corporate Reputation. Front. Psychol. 12:685351. doi: 10.3389/fpsyg.2021.685351

INTRODUCTION

In the Internet age, the huge demand from consumers for online travel has spawned online travel agencies (OTAs) that resemble a kind of consumption fashion (Pinto and Castro, 2019). The OTA Expedia was founded in 1996 (Ling et al., 2014), and it has become the online sales channel with the highest booking rate. The main function of the OTA is to obtain commission by providing travel related services, integrate products and reduce costs, and provide consumers with cheaper solutions (Kim et al., 2009). In essence, OTAs are intermediaries that rely on Internet platforms to provide online consulting, commenting, and booking services; because OTAs can easily compare products, prices, discounts, independent reviews, and photos of various hotels to meet the needs and preferences of consumers, they have developed rapidly, and some famous brands have appeared, such as Priceline, Expedia, and Ctrip.

OTAs have been developing rapidly globally, but they are facing the challenge of service failure. At present, OTA service failure events reported by the media are endless. For example, the "10.1" holiday in 2020 is the first holiday in China since its COVID-19 situation went under control. Some Chinese OTA brands, such as Ctrip, Qunar, and Tuniu, have become the object of consumer complaints. Mr. Han of Hangzhou, China, is a tourist who experienced failure in the OTA service. On October 1, 2020, he booked a reservation at a hotel in Sanya, China through the online platform of Ctrip, a famous OTA in China. On October 2, 2020, Mr. Han and his family arrived at the hotel and found that the hotel room was not as large as Ctrip had promised. The sanitation was poor, and it was not convenient to wash in the hotel room. Therefore, Mr. Han asked for a refund, but the hotel and Ctrip refused to give a full one, only 20%, and members of the hotel service staff were

arrogant. In the dispute, they attacked Mr. Han personally, causing his anger. Later, Mr. Han made a complaint through the Chinese consumer protection agency and exposed it through the media, which had a negative impact on the image of Ctrip and the hotel.

From the OTA case of Mr. Han, it can be seen that service failure has become a lingering curse for OTAs and service providers (such as hotels), which has created new challenges for the OTA service operation, so service recovery is urgently needed. Service failure is a kind of service that fails to meet customer expectations and that causes customer dissatisfaction (Peng and Jing, 2005; Nikbin et al., 2015). The current literature discusses the quality measurement system of online shopping service recovery (Liu and Li, 2017) and the relationship between online shopping service, service recovery, and customer satisfaction and loyalty (Jian and Ke, 2017). At present, academic circles have also studied the impact of word-of-mouth on the service recovery effect of travel agencies, and the service recovery effect is measured by customer satisfaction and customer loyalty (Pai et al., 2019). They developed a practical manual on service failure and service recovery for the travel industry (Inkson, 2019). However, the current literature still lacks research results on OTA service failure and service recovery, so in reality it cannot fully explain and guide the service failure and service recovery of OTAs, thus it needs to be further studied.

Perceived risk is a derivative concept of psychological research, and it refers to the uncertainty and risk perceived by consumers in the process of purchasing products or services (Jeon et al., 2020). The impact of perceived risk on negative emotions has been discussed in academic circles, such as in the studies conducted by Yoon and Lee (2014), Mi et al. (2019), and Zheng et al. (2019). It is generally believed that perceived risk will strengthen the negative emotions of customers. However, the current research still lacks relevant research results in the context of OTA service failure and service recovery. Corporate reputation is also operational. It can affect customer perceived value and then affect consumer intention and decision-making behavior (He et al., 2018). An OTA is a kind of intermediary that provides online consultation, comment, and reservation services based on an Internet platform, and it is also an online platform. With the rise in the number of Internet services, some researchers, such as Li (2014), Zhao and Wang (2012), and Olavarria-Jaraba et al. (2018), have studied the reputation of online platforms. In the service recovery scenario, whether the corporate reputation of OTAs will play a moderating role among some variables still lacks discussion.

The study uses a scenario experiment to conduct empirical analysis, and innovatively introduces perceived risk, negative emotions, and corporate reputation into the research on the impact mechanism of the service recovery effect of OTAs. Specifically, this study will explore the relationship between perceived risk and service recovery satisfaction, negative emotions and service recovery satisfaction, and service recovery satisfaction and customer loyalty, and verify the moderating role of corporate reputation between perceived risk and service recovery satisfaction. The study helps to expand the vision of service recovery research of OTAs, better explain the impact mechanism of service recovery effect of OTAs, improve the service recovery effect of OTAs, promote the healthy development of OTA business models, and bring more value to society and customers.

LITERATURE REVIEW AND RESEARCH HYPOTHESIS

Online Travel Agency

Rianthong et al. (2016) have suggested that the development of Internet technology has profoundly changed the way tourists book appointments, and online booking channels have developed greatly, including OTAs. The first OTA, Expedia, was established in 1996 (Ling et al., 2014). Since then, OTAs have developed rapidly all over the world. OTAs provide an effective platform for consumers to browse, purchase travel services, and share travel information through websites, mobile devices, apps, and call centers. An OTA is a tool for marketing, searching, and booking; it can earn tourism market share by managing hotel room reservations (Sheng, 2019). In essence, the OTAs are intermediaries that rely on an Internet platform to provide online consultation, comment, and reservation services. In the hotel and tourism industries, OTAs, through online distribution channels, play a key role in providing consumers with more attractive products (Kim and Lee, 2004). OTAs have profoundly changed the way consumers buy tourism products, and they have a great development potential. Traditional tourism enterprises have also strengthened cooperation with them to meet the needs of tourists (Sheng, 2019).

The OTAs help to increase the visibility of a hotel, thus increasing the interest and occupancy rate of tourists (Ling et al., 2015). OTAs have played an important role in building the reputation of a hotel. One of the factors contributing to the success of Weiganghui Hotel in Hong Kong was the close cooperation with OTAs at the beginning of its operations and the expansion of its promotion scope (Tony, 2013). However, some scholars have different views on this issue, believing that OTAs have little value but they have gained a lot of income that should belong to hotels (Green and Lomanno, 2012). For consumers, booking travel services through OTAs has several advantages: easy to use, low price, time-saving, comfortable, and diversified service products (Liu and Zhang, 2014; Hao et al., 2015; Pappas, 2017). However, there is still a lack of discussion on OTA service failure and service recovery, and it is urgent for the academic community to conduct in-depth research on this topic in order to promote the healthy development of OTA business models.

Service Failure and Service Recovery

Service failure refers to the when the service provided by a service enterprise fails to meet the minimum acceptable standard of customers, and fails to meet the requirements and expectations of customers, resulting in customer dissatisfaction (Chaouali et al., 2020). From the perspective of customer expectations, service failure is a kind of service that fails to meet customer expectations and a contact environment that causes customer dissatisfaction (Peng and Jing, 2005; Nikbin et al., 2015). Since the 1980s, service failure has attracted the attention of scholars.

Scholars have conducted extensive discussions on the definition, classification, causes, influencing factors, and consequences of service failure, enriching the research on service failure (Bitner et al., 1990; Maxham, 2001; Sven et al., 2015; Liu et al., 2019).

After the service failure of an enterprise, to repair its image and recover loss, service recovery is urgently needed. Service recovery is not only an action taken against service failure, it is also the hard work of an enterprise to make the service meet the expectation of a customer (Zhong et al., 2011). Some scholars have discussed service recovery strategy, and they think that it is necessary to confirm, evaluate, explain, apologize, and compensate for service failure, so as to improve the satisfaction of service recovery (Boshof, 1999). Heejung (2012) has suggested that service recovery should consider personality and preference, such as taking different service recovery strategies according to the education background, age, and preference of a customer. The influence mechanism of service recovery satisfaction is a research hotspot. The current literature focuses on the influence of customer misconduct, time perception, cultural differences, customer psychological contract violation, economic compensation, and emotional compensation on customer satisfaction after service recovery (Valenzuela and Cooksey, 2014; Albrecht et al., 2017).

Generally speaking, the current research on service failure and service recovery is relatively active, which lays a theoretical foundation for this study. However, the current literature still lacks OTA service failure and service recovery research results, so there is a need to continue in-depth study.

Perceived Risk and Negative Emotions

Perceived risk is a derivative concept of psychological research, and refers to the uncertainty and risks perceived by consumers in the process of purchasing products or services (Jeon et al., 2020). Perceived risk is subjective. No matter how big the actual risk of a product or service is, if it is not perceived by consumers, it will not affect their consumption behavior (Wang et al., 2018). Cui (2015) believes that perceived risk exists in six aspects: body, performance, property, time, society, and psychology. Sjöberg (2007) found that emotion does play an important role in perceived risk and related attitude, in which people found that interest in risk (positive emotion) is positively correlated with perceived risk. In the online shopping situation, because of the virtual characteristics of a network environment, the online shopping behavior of consumers often faces more perceived risks. Some scholars divide perceived risk into information, transaction, distribution, and after-sales (Wang, 2020), while others divide perceived risk into financial, functional, time, and privacy (Ren et al., 2019). For the research on influencing factors of perceived risk, Byun and Ha (2016) discussed the relationship between information usefulness, source credibility, perceived risk, impulsive purchase, and purchase intention in online shopping behavior. The research shows that information usefulness has a negative impact on perceived risk. Ren et al. (2019) suggested that in an online shopping scenario, the reputation, product quality, and website construction of a seller and other factors significantly negatively affect the financial, functional, and time risks in perceived risk, and that logistics support significantly negatively affects the financial, functional, time, and privacy risks in perceived risk.

Emotion is a psychological concept, and refers to the response of an individual to a specific object with external stimulation. Customer emotion is the synthesis of a series of related emotional reactions generated in the consumption process (Jia and Zhao, 2018). Previously, Westbrook and Oliver (1991) proposed that customer emotion can be divided into five dimensions: pleasant surprise, unpleasant surprise, anger, happiness, and sadness or indifference. However, most scholars divide emotions into positive and negative. Positive emotions include state of high energy activation, concentration, happiness, and engagement; negative emotions include anger, complaint, depression, regret, and helplessness (Berry et al., 2010; Jaeger et al., 2018). Liljander and Strandvik (1997) have suggested that customers will experience different positive and negative emotions on service, which affects customer satisfaction. Among them, negative emotions have a great impact on customer satisfaction. Ou and Verhoef (2017) have explored the incremental effects of positive and negative emotions on loyalty intention. The effects of these two emotions on loyalty intention are increasing, and positive emotions weaken positive connection (negative interaction). Khatoon and Rehman (2021) have stated that the negative emotions of consumers about a brand can be directly translated into actions against it, such as spreading negative word-of-mouth, avoiding, and retaliating.

The impact of perceived risk on negative emotions has been discussed in the current academic community. Zheng et al. (2019) conducted an empirical study on the Wenchuan earthquake-stricken areas in China and found that the perceived risk of residents in the disaster areas will enhance negative emotions. Some researchers have explored the impact of risk perception on subsequent behavior of apartment hotels. Studies show that perceived risk of apartment hotels has significant positive and negative effects on negative emotions and trust (Mi et al., 2019). Lee and Jung (2015) discussed the relationship between perceived risk and negative emotion in online shopping. The research shows that the perceived risk of customers will enhance their negative emotions. Yoon and Lee (2014) conducted an on-the-spot survey on tourists at the Seoul Lantern Festival. The results show that perceived risk has a positive impact on negative emotions. When the OTA service failure, such as being threatened by service personnel or unable to procure a refund occurs, customers will lose their spirit and money, and the perceived risk of customer dissatisfaction will arise. This will enhance the negative emotions of customers, including anger, complaint, depression, regret, and helplessness. Therefore, the study proposes the following hypothesis:

H1: Perceived risk has a significant positive impact on negative emotions.

Service Recovery Satisfaction

Since the 1960s, customer satisfaction has been a hot issue in service research. At present, there are two main perspectives to define customer satisfaction. The specific perspective of customer satisfaction is that customer satisfaction is a kind of

immediate emotional reflection of customer satisfaction with the value obtained after consumption in a specific situation; while customer satisfaction from an overall perspective is that customer satisfaction is a holistic attitude based on experience formed by customer attitude (Woodside et al., 1989). Combined with its characteristics, the study will adopt a specific perspective to define customer satisfaction. It is considered that service recovery satisfaction is a kind of feeling state in which the actual perceived effect of instant service recovery is higher than the expected service recovery. Its most prominent feature is a kind of customer satisfaction in a specific situation of service recovery, which is a kind of "second-degree satisfaction."

The relationship between perceived risk, customer satisfaction, and service recovery satisfaction has been studied in the academic field. Wang et al. (2017) took customers in the mobile communication industry of China as research samples, and the research confirmed that customer perceived risk would reduce customer satisfaction, and that customer participation would reduce customer risk perception. Yao and Deng (2017) suggested that the two dimensions of perceived risk include financial and time risks, and that these risks have a negative impact on customer satisfaction. Hsieh and Tsao (2014) pointed out that in a network environment, user perceived risk has a significant negative impact on user satisfaction, and that the higher the perceived risk, the lower the user satisfaction. Scridon et al. (2020) conducted an empirical study on the Romanian market, and the results showed that the perceived risk of customers has a significant negative impact on customer satisfaction and customer loyalty, and ultimately affects the profits of enterprises. Chang and Hsiao (2008) conducted an empirical study on the service recovery of hotels, and suggested that perceived risk will reduce customer value and that the decrease in customer value will inevitably reduce customer satisfaction. However, there is no research on the relationship between perceived risk and customer satisfaction in an OTA service recovery situation. In the OTA service recovery scenario, if the perceived risk of customers increases, the customer satisfaction reduces, and if the perceived risk of customers is reduced, the customer satisfaction is improved. Therefore, the following hypothesis is put forward:

H2: Perceived risk has a significant negative impact on service recovery satisfaction.

Westbrook (1991) studied the relationship between emotions and satisfaction. It is suggested that the positive emotions of customers have a positive effect on satisfaction, and that negative emotions have a negative effect on satisfaction. Yoon and Lee (2014) conducted an on-the-spot investigation on the tourists at the Lantern Festival in Seoul, South Korea, and analyzed them using a structural equation model. The results showed that perceived risk has a positive impact on negative emotions, and that negative emotions will reduce customer satisfaction. In the context of service recovery, Du and Fan (2007) explored the relationship between negative emotions and service recovery satisfaction using the method of simulation experiment. The results showed that the negative emotions of customers were negatively correlated with the service recovery satisfaction. In the OTA service recovery scenario, when the negative emotions of customers increase, the satisfaction of service recovery decreases. Therefore, the following hypothesis is put forward:

H3: Negative emotions have a significant negative impact on service recovery satisfaction.

Corporate Reputation

Reputation plays an important role in business. A good reputation can bring higher reputation benefits to enterprises, attract better employees, reduce costs, gain price advantages, and reduce enterprise risks (Dong, 2015). Corporate reputation also has operability. It can affect customer perceived value and then affect consumer intention and decision-making behavior, and it can help enterprises transform or obtain favorable factors from other aspects (He et al., 2018). With the rise of Internet services, some researchers have studied the reputation of online platforms. Olavarria-Jaraba et al. (2018) think that because of the lack of sound network supervision and social credit systems, if the default of a borrower is not curbed in time, the reputation risk of an online platform will increase. A good reputation is not only the precondition for P2P to charge higher service fees but also an effective incentive to improve the service level of fund lending and establish and maintain the reputation of an online platform (He et al., 2018; Olavarria-Jaraba et al., 2018).

In essence, an OTA is an intermediary organization that provides online consultation, comment, and reservation services based on an Internet platform, and it is also an online platform. Although the current literature on the impact mechanism of online platform service recovery effect is still very lacking, there is also a small amount of related research. Some studies suggest that online platform enterprises can become a collective signal of online platform reputation through social discussion of electronic word-of-mouth, thus driving customer demand, and that the establishment of online platform reputation can maintain long-term psychological contract relationship and reputation (Li, 2014). The credibility of online platform reputation information perceived by online platform buyers has a positive impact on the use level of reputation information and online shopping amount (Zhao and Wang, 2012). In the OTA service recovery scenario, if the OTA has a good corporate reputation, it will increase the sense of trust and satisfaction of a customer, which is helpful in enhancing service recovery. Therefore, the following hypothesis is put forward:

H4: Corporate reputation plays a positive moderating role in the relationship between perceived risk and service recovery satisfaction.

Customer Loyalty

Customer loyalty is the deep commitment of customers to their preferred enterprises or brands from which they will continue to buy repeatedly in the future. No matter how the situation or marketing power influences, customer loyalty will not produce switching behavior (Oliver, 1999). Loyal customers are the source of competitive advantage and an important guarantee of enterprise development. In the dimension research, customer loyalty can be divided into cognitive component, emotional component and behavior, reconstruction intention,

repeat purchase, recommendation to others, and attention (Li, 2014). In the online shopping environment, the academic research combines traditional customer loyalty driving factors and online shopping environment to study customer loyalty in online shopping. Luarn and Lin (2003) studied the driving factors of online shopping loyalty and found that customer trust, customer satisfaction, and customer perceived value have a positive impact on customer loyalty. Yu et al. (2008) took the traditional service industry as the research object, selected customer samples from Changsha, Hangzhou, Guangzhou, and seven other cities in China for empirical research, and found that service recovery satisfaction has a significant positive impact on customer loyalty. In the case of OTA service recovery, when customers are satisfied with OTA service recovery, they will inevitably produce behavioral loyalty and attitudinal loyalty. Therefore, the following hypothesis is put forward:

H5: Service recovery satisfaction has a significant positive impact on customer loyalty.

From the above analysis and hypothetical relationship, we can see that in the OTA service recovery scenario, negative emotions play a mediating effect between perceived risk and service recovery satisfaction. At the same time, perceived risk has an impact on customer loyalty through two variables: negative emotion and service recovery satisfaction. Negative emotion and service recovery satisfaction play a mediating effect between perceived risk and customer loyalty. Therefore, this study proposes two hypotheses for mediating effects:

H6: Service recovery satisfaction plays a mediating effect between perceived risk and customer loyalty.

H7: Negative emotions and service recovery satisfaction play a mediating effect between perceived risk and customer loyalty.

Based on the above seven hypotheses, a research model is proposed, as shown in **Figure 1**.

RESEARCH DESIGN

Experimental Design

The purpose of the study is to explore the influence mechanism of perceived risk and negative emotions on service recovery effect after OTA service failure. If the questionnaire survey method is used, only the respondents can be asked to recall the service failure and service recovery experience to fill in the relevant content, and the results may have a large memory bias (recall bias), affecting the accuracy of the research results. In contrast, the experimental method based on scenario description has higher internal validity. Combined with the characteristics of this study, based on the experimental research methods of Du and Fan (2007) and Poggi and D'Errico (2010), we decided to use the scenario experiment method for empirical research.

In order to improve the external validity of the situational experiment, the materials of the situational experiment are obtained from the OTA service failure events reported by the Chinese media. At the same time, this study has also referred to the online tourism complaints published by the China National Tourism Administration, and has made an appropriate adaptation to make it more suitable for scenario experimental research.

The service recovery scenario is designed as 2 (high perceived risk vs. low perceived risk) × 2 (high negative emotions vs. low negative emotions) × 2 (high corporate reputation vs. low corporate reputation). To avoid the mixed effect caused by the existing relationship between customers and real OTAs, the OTA brand, hotel name, and travel route name in the experimental situation are fictitious. The experiment was divided into three parts.

The first part describes the perceived risk of customers. (1) Under the experimental conditions of high perceived risk, the study describes that the location of the hotel where the OTA arranges customers to stay is not good, and that the room facilities and sanitary conditions are also relatively poor, which is inconsistent with the introduction on the OTA platform. Customers ask for a refund, but the OTA and the hotel refuse to refund. In negotiation, the service attitude of the hotel service staff was also poor, and they attacked the customers with language. Subsequently, the customers complained to the OTA and the hotel. (2) Under the experimental conditions of low perceived risk, this paper describes that the OTA arranges customers to stay in a hotel. The room facilities of the hotel are basically consistent with the introduction on the OTA platform, but the customers are not very satisfied with the hotel location and sanitation, and they think the price is relatively high. The customers have not complained about the OTA.

The second part describes the negative emotions of customers. (1) Under the experimental conditions of high negative emotions, the hotel location, room facilities, and sanitary conditions arranged by the OTA were poor, and the service attitude of the service staff was also poor, and the staff even made verbal attacks on customers. The customers feel very angry, frustrated, regretful, and helpless, and the negative emotions are very obvious. (2) Under the experimental condition of low negative emotion, this study describes that after the OTA arranges customers to stay in the hotel, although the room facilities are basically consistent with the introduction on the OTA platform, the customers are not satisfied with the hotel location, room hygiene, and other conditions. On the whole, the degree of negative emotions of the customers is not so high, just slight negative emotions, such as slight complaints and regrets about service details, which are attributed to bad luck. The customers did not show serious anger, depression, and other negative emotions.

The third part describes the corporate reputation of an OTA. (1) Under the experimental conditions of high corporate reputation, the OTA enjoys a high level of reputation, and its service quality has a good reputation, which is welcomed by customers and the public. (2) Under the experimental condition of low corporate reputation, the study describes that the OTA has low popularity and reputation, and that its word-of-mouth is also relatively poor. There are frequent reports of OTA service failure in the media, and it is not welcomed by customers and the public.

Since in the scenario experiment the service recovery scenario is designed as 2 (high perceived risk vs. low perceived risk) ×

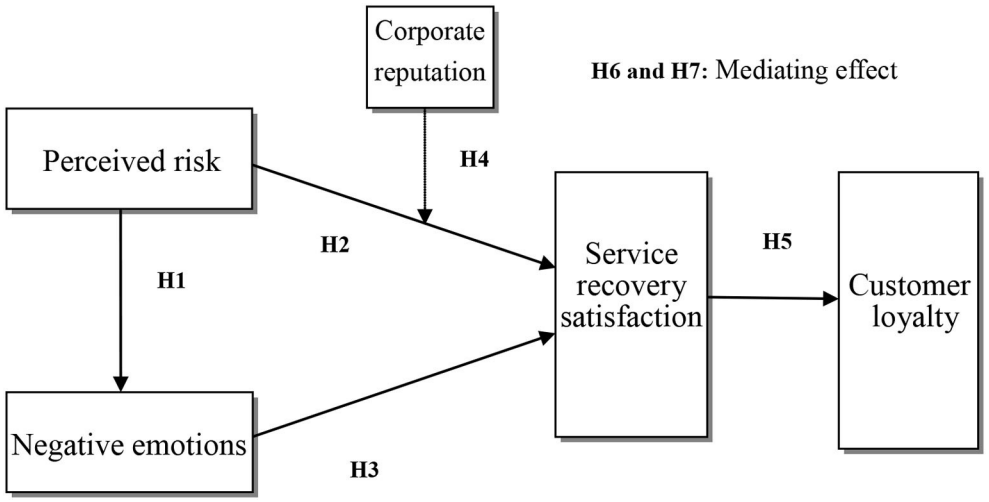


FIGURE 1 | Research model.

2 (high negative emotions vs. low negative emotions) × 2 (high corporate reputation vs. low corporate reputation), there are eight scenarios in total. The researcher asked the subjects to put themselves in a certain situation, explained various scenarios, and guided the subjects to imagine their feelings and reactions as OTA customers in the service failure and service recovery scenarios of the OTA. Then, the researcher instructed the subjects to fill in the questionnaire of the situational experiment.

Variable Measurement

The measurement variables include perceived risk, negative emotions, corporate reputation, service recovery satisfaction, and customer loyalty. In order to ensure the pertinence and effectiveness of measurement of the five variables, this study adopts the relevant scales of authoritative literature and makes corresponding modifications according to the OTA service recovery scenarios.

In the study, the measurement of perceived risk draws on the scales of Li and Li (2007) and Lee and Moon (2015), and eight items are set. In the development of the scale of negative emotions, seven Chinese residents with OTA service failure and recovery experience were invited to speak about the related negative emotions in words in the OTAs service recovery situation, and then the frequency of these words was calculated and sorted to form the initial negative emotion scale. On this basis, the study referred to the emotion measurement scale used by Baron et al. (2005) and Fang et al. (2019), and finally set five items for negative emotion variables. Based on the research results of Ponzi et al. (2011), and Wang and Guo (2018), the corporate reputation scale has five items. According to Bhatta and Premkumar (2004) and Ribbink et al. (2004), the measurement of service recovery satisfaction is modified according to the characteristics of OTA service failure and service recovery. The measurement of customer loyalty is based on the customer loyalty scale compiled by Zeithaml et al. (1996), and four items are set up, including two aspects of customer behavior loyalty and customer attitude loyalty. On this basis, the study also invited 15 Chinese citizens with the OTA service experience, and four psychology and management researchers to revise the text content and expression of the first draft of the questionnaire, and correct the ambiguous meaning and unclear expression in the questionnaire. All of the items were measured using a five-point Likert scale with the options "very consistent," "consistent," "basically consistent," "not very consistent," and "not consistent."

Pre-experiment

Before the implementation of the formal experiment, to ensure the reliability of the experimental situation and the effectiveness of experimental manipulation, the researchers conducted a pre-experiment on college students. The college students participated in the pre-experiment as volunteers. The reason why the pre-experiment chooses the college students as experimental samples is that the internal consistency of the college students is high, and the college students are curious and easy to accept new things. They are the active group of OTA consumption, and they have a better understanding of the OTA operation and consumption process, which has a good experimental value.

Eighty-two college students, 39 females (47.6%) and 43 males (52.4%), who had OTA consumption experience participated in the pre-experiment. Independent sample *T*-test was conducted on the questionnaire data. The test results show that the credibility of the experimental scenario is relatively high (the average is 3.83), and that the evaluation of perceived risk, negative emotions, and corporate reputation of the subjects in different experimental scenarios is consistent with the experimental design. Under the conditions of high perceived risk, the perceived risk of OTA service failure was higher ($M_{\text{high perceived risk}} = 4.46$, $M_{\text{low perceived risk}} = 2.97$; $T = 2.943$, $DF = 54$, $P < 0.01$). Under the experimental conditions of high negative emotion, the negative emotion of OTA service failure was stronger ($M_{\text{high negative emotions}} = 4.48$, $M_{\text{low negative emotions}} = 3.29$, $T = 2.819$, $DF = 73$, $P < 0.01$). Under the experimental conditions

of high corporate reputation, the subjects rated the corporate reputation of the OTA better ($M_{\text{high corporate reputation}} = 4.62$, $M_{\text{low corporate reputation}} = 3.37$, $T = 2.311$, $DF = 62$, $P < 0.05$). The independent sample T-test results of the above pre-experiment show that the variables of perceived risk, negative emotions, and corporate reputation are manipulated successfully.

Formal Experiment

From October to November 2020, the study conducted a formal experiment. The experimental samples were from Guilin, Nanning, and Liuzhou, and farmers from the Yangshuo and Luzhai counties in China. The formal experiment was composed of researchers, four university volunteers, and 12 basic cadres of communities and villages. The subjects were recruited in the formal experiment through the following two ways. The first was through cooperation with the communities in Guilin, Nanning, and Liuzhou (community is grassroots management organizations in Chinese city) to collect experimental samples through social platforms (QQ group and WeChat group). The subjects were required to have OTA service purchase experience. The second was through cooperation with the villages of the Yangshuo and Luzhai counties (village is grassroots management organization in Chinese rural areas) and collection of subjects through the social platform (QQ group). The subjects were also required to have OTA service purchase experience.

Because the subjects lived in different areas, the formal experiment was completed 11 times. The service recovery scenario designed in this study is 2 (high perceived risk vs. low perceived risk) × 2 (high negative emotions vs. low negative emotions) × 2 (high corporate reputation vs. low corporate reputation). In each experiment, the researcher explained various situations, asked the subjects to choose one of the scenarios, read the text description of their own situation carefully, and put the individual under the situation. The researcher guided the subjects to imagine their feelings and reactions as OTA customers in a certain situation of OTA service failure and service recovery. Then, the researcher instructed the subjects to fill in the questionnaire of the situational experiment. To improve the quality of the questionnaire, the researchers provided necessary guidance. At the same time, in order to ensure the filling effect and respect personal privacy, the subjects were asked to complete the answers in free private space.

Formal experiments focus on solving the problem of social desirability bias. Scholars believe that there are three main reasons for the occurrence of social desirability bias: the subjects themselves have higher social desirability, the test situations stimulate the social desirability bias of the subjects, and the items themselves cause social desirability bias (Paulhus et al., 2002). In this experiment, the subjects involved ordinary Chinese citizens and farmers who did not have high social desirability and would make a more objective evaluation based on facts. There is no social desirability in the guidance of the questionnaire. We have also optimized the expression of items to prevent the occurrence of social desirability bias. Before the test, we told the subjects that the questionnaire was anonymous and would not reveal privacy, so that they could fill in the questionnaire objectively and safely.

In the process of investigation, there are many factors that lead to the missing or non-response of some individuals in the sample, so the final valid sample is only a part of the survey sample (Smironva et al., 2019). Therefore, formal experiments also focus on solving the problem of non-response bias. First of all, when the subjects are recruited through community and village committees, they are required to choose to participate in the experiment voluntarily and have the OTAs service purchase experience, which greatly reduces non-response bias. Second, in the formal experiment, we trained the subjects and informed them of the experimental operation process and questionnaire filling and answering methods, which improved the accuracy of the questionnaire and avoided the question of non-response bias.

A total of 388 questionnaires were collected in the formal experiment, and 355 questionnaires were valid. The effective rate was 91.26%. The sample distribution is shown in **Table 1**.

DATA ANALYSIS

Reliability and Validity Test

Cronbach's α is an index used to measure the reliability of the questionnaire data. Cronbach's α is required to be greater than the threshold value of 0.7. The higher the Cronbach's α is, the stronger the internal consistency of the questionnaire is and the higher the reliability is (Nunnally and Bemstein, 1994; Hair and Black, 2006). As shown in **Table 2**, the Cronbach's α of the five variables ranged from 0.741 to 0.882. If the Cronbach's α is >0.7, it means that reliability has passed the test (Hair and Black, 2006), indicating that the internal consistency of the questionnaire met the requirements, and that the reliability of the questionnaire passed the test.

Based on reliability analysis, the validity of the questionnaire is analyzed. In terms of content validity, all items of the questionnaire refer to published authoritative journals, widely refer to the opinions of Chinese residents with the OTA service failure and recovery experience, and make adjustments combined with the OTA service recovery scenarios, which shows that the questionnaire items have good content validity. Convergence validity is shown in **Table 2**. The standardized load factors of all items are >0.5, the T-value is greater than the threshold value of 1.96, the combined reliability (CR) values of the five variables are >0.7, and the average extraction variance (AVE) is >0.5, which meets the convergence validity test standard of Hair and Black (2006) and Wu (2010). Therefore, the questionnaire passed the convergence validity test.

The analysis of discriminant validity is shown in **Table 3**. The square root value of the AVE of the five variables is greater than the correlation coefficient between the variable and other variables. According to Wu (2010) criterion of discriminant validity, the questionnaire has good discriminant validity, and the discriminant validity has passed the test.

In the test for construction validity, the measurement values of the research model fit are shown in **Table 4**: $\chi^2/df = 2.752$, CFI = 0.945, TLI = 0.942, SRMR = 0.039, RMSEA = 0.058. According to the good model standard of Wu (2010), χ^2/df should be <5, CFI and TLI should be >0.9, SRMR should be <0.05, and RMSEA should be <0.1. Therefore, the above indicators of the research model have reached the standard of a good model, and the construction validity of the questionnaire has passed the test.

TABLE 1 | Sample distribution.

One-Class indicators	Two-Class indicators	Sample size	Percentage	One-Class indicators	Two-Class indicators	Sample size	Percentage
Gender	Male	172	48.45%	Education	Secondary school and below	121	34.08%
	Female	183	51.55%		College degree	123	34.65%
Region	Guilin city	104	29.30%		Bachelor	78	21.97%
	Nanning city	86	24.23%		Master and doctor	32	9.01%
	Liuzhou city	77	21.7%	Occupation	Enterprise staff	74	20.85%
	Yangshuo county	47	13.24%		Professional	61	17.18%
	Luzhai county	41	11.55%		Self-employed person	54	15.21%
Age	Under 25 years	95	26.76%		Civil service staff	23	6.48%
	26–35 years	103	29.01%		Student	42	11.83%
	36–59 years	127	35.77%		Farmer	68	19.15%
	Over 60 years	30	8.45%		Other	33	9.3%

TABLE 2 | Test for reliability and convergent validity.

Variables	Items	Normalized load factor	*T*-value	Cronbach's α	CR	AVE
Perceived risk	1. The OTA may have forged its credit record	0.783	3.481	0.882	0.921	0.597
	2. The complaint was not handled in time	0.761	4.827			
	3. My payment cannot be returned	0.691	2.959			
	4. Poor attitude of service staff	0.903	4.724			
	5. I am worried about wasting a lot of time	0.731	3.765			
	6. I am worried about the abuse of personal information	0.843	5.236			
	7. People around me may have a lower evaluation of me	0.789	6.382			
	8. The service failure made me feel bad	0.649	2.168			
Negative emotions	9. The service failure disappointed me	0.812	2.996	0.769	0.863	0.560
	10. Service recovery did not eliminate my complaint	0.636	3.289			
	11. I am angry about the OTA	0.783	6.032			
	12. I feel angry about the hotel	0.824	3.664			
	13. I regret choosing the OTA	0.666	3.072			
Corporate reputation	14. The OTA is trustworthy	0.905	6.137	0.857	0.902	0.650
	15. People around me have recommended the OTA to me	0.893	7.195			
	16. The OTA has a good reputation	0.782	5.396			
	17. There are few negative reports about the OTA	0.728	4.214			
	18. The OTA is a good partner in my life	0.673	2.797			
Service recovery satisfaction	19. I am satisfied with the timing of service recovery	0. 677	2.825	0.741	0.839	0.513
	20. I am satisfied with the way of service recovery	0.779	2.821			
	21. Service recovery solves my problem	0.833	5.264			
	22. Service recovery meets my psychological expectation	0.626	6.259			
	23. I am satisfied with the service staff	0.643	3.998			
Customer loyalty	24. I will continue to purchase the services of the OTA	0.872	7.223	0.809	0.848	0.585
	25. I will recommend others to use the OTA's services	0.739	3.262			
	26. I will be a fan of the OTA	0.801	7.298			
	27. The OTA gives me a sense of belonging	0.627	4.465			

Test for Direct Impact Relationship

In this study, the SPSS22.0 software was used for multilevel regression analysis to test direct impact relationship and moderating role. In order to avoid multicollinearity, independent variables and regulatory variables were analyzed and treated centrally before multi-level regression analysis. The calculation results of the variance expansion factor (VIF) show that the VIF is between 2.072 and 3.382, which is lower than the empirical value of 10, indicating that there is no multicollinearity problem in the research model. The results of multilevel regression are shown in **Table 5**.

This study constructs models 1 and 2 to test the direct relationship between perceived risk and negative emotions. As shown in **Table 5**, model 1 is the regression of negative emotions

TABLE 3 | Test for discriminant validity.

Variables	1	2	3	4	5
1. Perceived risk	0.773				
2. Negative emotions	0.597	0.776			
3. Corporate reputation	−0.194	−0.186	0.806		
4.Service recovery satisfaction	−0.358	−0.322	0.412	0.716	
5. Customer loyalty	−0.213	−0.303	0.314	0.589	0.765

The value on the diagonal is the square root of AVE; and the other data are the correlation coefficients of the corresponding variables.

TABLE 4 | Research model fit.

Fit index	χ2	*df*	χ^2/df	CFI	TLI	SRMR	RMSEA
Index value	308.215	112	2.752	0.945	0.942	0.039	0.058

with control variables. The results show that there is a weak negative relationship between age and negative emotions, and that the relationship between gender, origin, education, and perceived risk is not statistically significant. Model 2 showed that perceived risk had a significant positive impact on negative emotions ($\beta = 0.697$, $P < 0.01$). At the same time, both models 1 and 2 passed the F-test, and the ΔR^2 of model 2 was >0, indicating that the explanatory power of the model gradually increased. Therefore, hypothesis H1 passed the test.

The direct impact of perceived risk and negative emotion on service recovery satisfaction was tested by constructing models 3, 4, and 5. First, the control variables are included in model 3, which shows that there is a weak negative impact between education and perceived risk, and that the impact of gender, age, source region, and other control variables on perceived risk is not statistically significant. Second, according to the method of layer-by-layer inclusion, perceived risk is included in model 4. The results show that perceived risk has a significant negative impact on service recovery satisfaction ($\beta = -0.589$, $P < 0.05$), so research hypothesis H2 passes the test. On this basis, negative emotions were included in model 5. The results showed that negative emotions had a significant negative impact on service recovery satisfaction ($\beta = -0.665$, $P < 0.01$). The results supported hypothesis H3.

To test the relationship between service recovery satisfaction and customer loyalty, the study constructs models 7 and 8. First, the control variables are included in Model 7, which shows that the negative influence of education, age, occupation, and other control variables on customer loyalty is not statistically significant. Then, service recovery satisfaction is included in model 8. The results show that service recovery satisfaction has a significant impact on customer loyalty ($\beta = 0.708$, $P < 0.01$). The results support research hypothesis H5.

Moderating Role Test

In this study, moderating role was also tested by multilevel regression analysis. As shown in **Table 5**, in model 6, the moderating role of corporate reputation between perceived risk and service recovery satisfaction is tested. From the multilevel regression results, it is found that the interaction coefficient between perceived risk and corporate reputation is positive and statistically significant ($\beta = 0.217$, $P < 0.05$), indicating that in the OTA service recovery, corporate reputation plays a positive moderating role in the relationship between perceived risk and service recovery satisfaction, and that research hypothesis H4 is supported.

To more clearly show the moderating role of corporate reputation between perceived risk and service recovery satisfaction, this study draws a moderating role diagram, as shown in **Figure 2**. The results of simple slope test show that perceived risk has a strong negative impact on service recovery satisfaction when corporate reputation is low ($\beta = -0.728$, $P < 0.05$). When corporate reputation is high, perceived risk has a weak negative impact on service recovery satisfaction ($\beta = -0.316$, $P < 0.05$). Therefore, when corporate reputation reduces the negative impact of perceived risk on service recovery satisfaction, it plays a positive moderating role. Research hypothesis H4 has been further verified.

Mediating Effect Test

Because the mediating effect test in the study belongs to simple mediating effect test, and this kind of mediating test using Bootstrap method has obvious advantages, which will be more scientific and accurate than the causal stepwise regression method (Preacher et al., 2007). At the same time, the sampling performed in this study does not conform to the normal distribution, and the Bootstrap method does not need to assume normal distribution of sampling. It estimates the indirect effect and sampling distribution through repeated sampling, and estimates the confidence interval of indirect effect according to distribution characteristics (Preacher et al., 2007; Wu, 2010). Therefore, this study will use the Bootstrap method for mediating effect test.

The study constructs the competition model and chain mediating effect model to further confirm the existence of chain mediating effect. **Table 6** shows the fit index of the competition model and the chain mediating effect model. According to Wu (2010), the standard of good model is that χ^2/df is <5, CFI and TLI are >0.9, SRMR is <0.05, and RMSEA is <0.1. As shown in **Table 6**, the χ^2/df value of the competition model is 6.363, and other fitting indexes do not meet the good requirements, indicating that the fitting degree of the competition model is poor. The χ^2/df, CFI, TLI, SRMR, and RMSEA fit indexes of the chain mediating effect model all meet the requirements of a good model and are significantly better than the competition model parameters. Therefore, the chain mediating effect does exist.

The mediating effect test is shown in **Table 7**. The indirect effect value of mediating path "perceived risk → service recovery satisfaction → customer loyalty" is 0.276, accounting for 46.7% of the total effect. It shows that service recovery satisfaction plays a mediating effect between perceived risk and customer loyalty. Research hypothesis H6 passes the test. The indirect effect value of mediating path "perceived risk → negative emotion → service recovery satisfaction → customer loyalty" is 0.132, accounting for 22.33% of the total effect, which indicates

TABLE 5 | Multilevel regression analysis.

Variables	Negative emotions		Service recovery satisfaction				Customer loyalty	
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
Intercept	3.032* (0.044)	2.839** (0.032)	3.566* (0.027)	2.681* (0.089)	2.929* (0.075)	3.528* (0.109)	2.434** (0.046)	3.633** (0.113)
Control variables								
Gender	0.091	−0.022	0.091	0.014	0.011	0.122	0.023	0.044
Region	0.012	0.017	0.018	0.027	0.008	0.031	0.021	0.023
Age	−0.103*	−0.096	−0.073	−0.241	−0.111	−0.019	−0.099	−0.066
Education	0.028	0.024	−0.043*	0.208*	0.016	0.035*	0.031	0.011*
Occupation	0.018	0.013	−0.012	0.102	0.021	0.025	0.017	0.039*
Independent variable								
Perceived risk		0.697**		−0.589*		0.499*		
Negative emotions					−0.665**			
Service recovery satisfaction								0.708**
Moderating variables								
Corporate reputation						0.318*		
Interaction item								
Perceived risk × corporate reputation						0.217*		
R^2	0.041	0.242	0.053	0.137	0.324	0.302	0.049	0.132
ΔR^2		0.073		0.102	0.239	0.083		0.225
F	4.203**	2.807**	3.279**	1.986*	3.229**	2.906**	2.154*	2.652**

*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

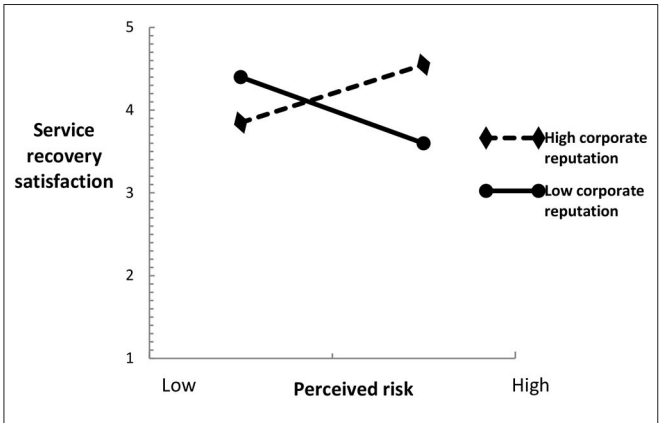


FIGURE 2 | Moderating role chart.

TABLE 6 | Test of competition model of mediating effect.

Model	χ^2	*df*	χ^2/df	CFI	TLI	SRMR	RMSEA
Chain mediating effect	57.793	21	2.752	0.934	0.925	0.037	0.054
Competition model	178.176	28	6.363	0.711	0.802	0.074	0.123

that negative emotion and service recovery satisfaction play a mediating effect between perceived risk and customer loyalty. Research hypothesis H7 passes the test. As shown in **Table 7**, the confidence intervals of the total mediating effect and the two mediating effects do not contain a value of 0, and are statistically significant. The total mediating effect was 0.408, accounting for 69.03% of the total effect.

CONCLUSION AND DISCUSSION

Conclusions and Theoretical Contributions

First, the results show that: in the OTA service recovery scenario, the perceived risk of customers has a significant positive impact on negative emotions, and that the perceived risk and negative emotions of customers have a significant negative impact on service recovery satisfaction. Therefore, the above results support hypotheses H1, H2, and H3. H1 is supported, which shows that in the OTA service recovery scenario, the greater the perceived risk, the greater the negative emotion. This research conclusion extends the research scenario to that of OTA service recovery, which is an extension and enrichment of the conclusions of Lee and Jung (2015) and Mi et al. (2019). This study verifies hypothesis H2, which confirms and extends the conclusions of Hsieh and Tsao (2014), Chang and Hsiao (2008), and Scridon et al. (2020) that perceived risk has a negative impact on customer satisfaction and service recovery satisfaction. This study confirms that the perceived risk of customers will have a negative impact on service recovery satisfaction in the OTA service recovery scenario. At the same time, this study also verifies research hypothesis H3, which indicates that negative emotions of customers will also have a negative impact on service recovery satisfaction. In previous studies, Du and Fan (2007) and Yoon and Lee (2014) have suggested that negative emotions will negatively impact customer satisfaction and service recovery satisfaction. The conclusion of this study is an extension of

TABLE 7 | Test results of mediating effect.

Mediating effect path	Indirect effect value	Standard error	Upper limit	Lower limit	Effect proportion
1. Perceived risk → service recovery satisfaction → customer loyalty	0.276	0.019	0.164	0.468	46.7%
2. Perceived risk → negative emotion → service recovery satisfaction → customer loyalty	0.132	0.008	0.123	0.391	22.33%
4. Total mediating effect	0.408	0.027	0.327	0.702	69.03%
5. Total effect	0.591	0.036	0.214	0.575	100%

previous studies. In short, through empirical research, research hypotheses H1, H2, and H3 are supported, which expand the connotation and applicable scenarios of the relationship between perceived risk, negative emotions, and service recovery satisfaction, and expand the connotation and applicable scenarios of the relationship between negative emotions and service recovery satisfaction.

Second, according to the results of this study, corporate reputation plays a positive moderating role in the relationship between perceived risk and service recovery satisfaction. Therefore, hypothesis H4 is supported. In the OTA service recovery scenario, with the improvement in OTA corporate reputation, service recovery satisfaction will increase faster. The corporate reputation of an OTA will positively moderate the relationship between perceived risk and service recovery satisfaction, that is, reduce the negative impact of perceived risk on service recovery satisfaction. The existing literature suggests that the corporate reputation of an online platform can maintain long-term psychological contract relationship, create better relationship, and improve demand (Li, 2014). However, there is insufficient research on the moderating role of corporate reputation between perceived risk and service recovery satisfaction under the OTA service recovery scenario. This study verifies that the corporate reputation of an OTA can positively moderate the negative impact of perceived risk on service recovery satisfaction, which makes up for the lack of current research. This study provides a theoretical and empirical explanation to further explore the regulatory mechanism of the corporate reputation of OTAs and is helpful for the majority of service enterprises, especially OTAs, in improving their reputation or re-evaluating their image.

Third, in the OTA service recovery scenario, service recovery satisfaction has a positive impact on customer loyalty, and research hypothesis H5 has passed the hypothesis test. Negative emotions and service recovery satisfaction play a mediating effect between perceived risk and customer loyalty. Hypothesis H6 passes the hypothesis test. The research also shows that service recovery satisfaction also plays a mediating role between perceived risk and customer loyalty, and that research hypothesis H7 also passes the hypothesis test. The conclusion that service recovery satisfaction has a positive impact on customer loyalty further confirms the research results of Yu et al. (2008), extends the research scenario to that of OTA service recovery, and expands the applicable scenarios and research connotation of service recovery satisfaction and customer loyalty. However, in the OTA service recovery scenario, there are few research studies on the mediating effects of negative emotions and service recovery satisfaction between perceived risk and customer loyalty. This study makes a breakthrough and innovation in this aspect.

Practical Applications

First, after service failure, OTAs and their service providers should manage the negative emotions of customers and optimize service recovery strategies. Because of service failure, the perceived risk of customers to OTAs and service providers will increase, and negative emotions, such as complaints, disappointment, and helplessness will follow. This study shows that perceived risk and negative emotions have a negative impact on service recovery satisfaction. Therefore, OTAs and service providers should manage and guide emotions of customers through the negative emotion management scheme designed by psychological and service management experts, enriching service recovery means, combining material compensation with spiritual compensation, timely carrying out personalized service recovery, taking interests of customers into consideration, and improving service recovery satisfaction.

Second, OTAs should strive to maintain and enhance their corporate reputation and do everything possible for the sake of customers to improve corporate reputation. The results show that corporate reputation reduces the negative impact of perceived risk on service recovery satisfaction. If an OTA has a good corporate reputation, in the process of service failure and service recovery, the negative effect of perceived risk on service recovery satisfaction will be reduced, and it will become the "protector" of enterprises to resist risks. However, for an OTA, a good corporate reputation cannot be formed overnight. It needs a long-term accumulation of OTA, which is the embodiment of corporate image. In the era of the Internet, OTAs should be honest, put an end to cheating customers, do everything possible for customers, and establish a good brand image. At the same time, OTAs need to carry out public relations and promotion work, and use news media, hot events, and public welfare activities to improve corporate reputation.

Third, OTAs should improve customer loyalty through effective service recovery. After OTA service failure, customers are very disappointed with OTAs, but they can reverse their attitude through service recovery. From the perspective of this study, reducing the perceived risk and negative emotions of customers is only an important incentive to enhance customer

loyalty, not a direct factor. Therefore, in OTA service recovery, we should pay attention to the formation of service recovery satisfaction. OTAs should take scientific and effective service recovery measures to reduce perceived risks and negative emotions of customers, scientifically guide customers, be good at approaching customers, resolve negative emotions of customers, solve concerns of customers about money and time, and improve service recovery satisfaction. On this basis, customers will continue to patronize the OTA and recommend it to their relatives and friends, so as to form customer loyalty and become loyal customers of the OTA.

Research Limitations and Prospects

First, future research needs to extend the research framework. The research framework of this study does not involve customer trust, consumer value, service quality, consumer perception, and other aspects, but in the process of service recovery, the above four variables may have an important impact on the effect of service recovery. In future research, we will extend the research framework by including consumer trust, consumer value, service quality, consumer perception, and other variables in order to better explore the impact mechanism of OTA service recovery effect. For example, we will study the mediating role of consumer trust between service recovery satisfaction and customer loyalty, or the mediating role of consumer trust between consumer forgiveness and customer loyalty. The purpose of the study is to examine the impact of consumer values on consumer emotion and service recovery satisfaction, find out the differences in the impact of consumers with different values on related variables, explore the impact of service quality on the relationship between negative emotions and service recovery satisfaction, and research on the relationship between consumer perception and emotion type, service quality and service recovery satisfaction, etc. In addition, in future research, this study can also go beyond the current research framework to explore if consumer distrust will lead to disapproval of OTAs and cause damage to the image and reputation of OTAs. These research ideas will expand the vision of OTA service recovery research and enrich the accumulation of service recovery theory.

Second, future research needs to optimize research methods. Although the scenario experiment is suitable for collecting customer data and has high internal validity, these data are all from the subjective evaluation of customers and have certain limitations. In future research, we will collect objective data of customers at the same time. Objective data mainly refer to the data generated by customers using OTA platforms, and they include the login time and times of customers on OTA platforms, length of stay of customers on OTA platforms, proportion of customers choosing various types of hotels on OTA platforms, and so on. Obtaining objective data will make up for the deficiency in scenario experiment, because subjective evaluation results of customers often have certain deviation. Therefore, the combination of objective and subjective data will be the direction of future research.

Third, future research needs to expand the source of samples. The samples in the study come from Guilin, Nanning, and Liuzhou in China, but there are no samples from other countries, especially from Southeast Asia. In future research, we will expand the scope of sample sources, investigate samples from many countries, expand sample size, and improve representativeness of samples. At the same time, adaptive joint analysis (ACA) embedded in polyhedral methods will be used to reduce the complexity of the questionnaire.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

This research is approved by the Ethics Committee of Guilin University of technology, which belongs to Guilin University of technology. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

JW independently completed the paper, carried out literature analysis, scenario experiment and data analysis, and demonstrated the research hypothesis, as well as other work in the paper.

FUNDING

Project supported by the National Natural Science Foundation of China (Grant No. 71962005). The effects of knowledge embeddedness in online review on product performance: A perspective of semantic association.

REFERENCES

Albrecht, A. K., Walsh, G., and Beatty, S. E. (2017). Perceptions of group versus individual service failures and their effects on customer outcomes: the role of attributions and customer entitlement. *J. Serv. Res.* 20, 188–203. doi: 10.1177/1094670516675416

Baron, S., Schoefer, K., and Ennew, C. (2005). The impact of perceived justice on consumers' emotional responses to service complaint experiences. *J. Serv. Market.* 19, 261–270. doi: 10.1108/08876040510609880

Berry, R., Tanford, S., Montgomery, R., and Green, A. J. (2010). How we complain: the effect of personality on consumer complaint channels. *J. Hosp. Tour. Res.* 42, 74–101. doi: 10.1177/1096348014550921

Bhatta, C., and Premkumar, G. (2004). Understanding changes in belief and attitude toward information technology usage: a theoretical model and longitudinal test. *Mis Q.* 28, 229–254. doi: 10.2307/25148634

Bitner, M. J., Booms, B. M., and Tetreault, M. S. (1990). The service encounter: diagnosing favorable and unfavorable incidents. *J. Market.* 54, 71–85. doi: 10.1177/002224299005400105

Boshof, C. (1999). An instrument to measure satisfaction with transaction specific service recovery. *J. Serv. Res.* 1, 65–74. doi: 10.1177/109467059913005

Byun, C. G., and Ha, H. (2016). The effects of usefulness and credibility of word-of -mouth information on purchase intention in social commerce: focused on the mediating roles of perceived risk and impulse buying. *J. Korea Serv. Manage. Soc.* 17, 89–106. doi: 10.15706/jksms.2016.17.4.005

Chang, H. S., and Hsiao, H. L. (2008).Examining the casual relationship among service recovery, perceived justice, perceived risk, and customer value in the hotel industry. *Serv. Indust. J.* 28, 513–528. doi: 10.1080/02642060801917646

Chaouali, W., Souiden, N., and Ringle, C. M. (2020). Elderly customers' reactions to service failures: the role of future time perspective, wisdom and emotional intelligence. *J. Serv. Market.* 35. doi: 10.1108/JSM-08-2019-0318

Cui, B. J. (2015). "Review and prospect of consumer perceived risk in the context of product injury crisis," in *Business Economics and Management*, 65–73.

Dong, X. Z. (2015). Research on online customer loyalty of platform shopping website: a double intermediary path based on relationship quality and website attachment. *J. Market. Sci.* 11, 104–128.

Du, J. G., and Fan, X. C. (2007). The effect of emotion on customer satisfaction and behavior after service recovery: a study based on the perspective of emotional infection. *Manage. World* 85–94.

Fang, S. J., Li, Y. Q., and Fu, Y. X. (2019). Symbolic or utilitarian? Research on recovery strategies of scenic areas based on emotional contagion theory. *Tour. Tribune* 34, 44–57.

Green, C. E., and Lomanno, M. V. (2012). *An AH&LA and STR Special Report. Distribution Channel Analysis: A Guide for Hotels.* McLean, VA: HSMAI Foundation.

Hair, J. F., and Black, B. (2006). *Multivariate Data Analysis.* Upper Saddle River, NJ: Prentice Hall Press, 113–121.

Hao, J., Yu, Y., Law, R., and Fong, D. (2015). A genetic algorithm-based learning approach to understand customer satisfaction with OTA websites. *Tour. Manage.* 48, 231–241. doi: 10.1016/j.tourman.2014.11.009

He, X. H., Hu, D. B., and Su, L. J. (2018). Perceived fairness, relationship quality and environmental responsibility behavior of residents in tourist destinations. *J. Tour.* 33, 117–131.

Heejung, R. O. (2012). Service recovery evaluations: GLBT versus heter customers. *Int. J. Hosp. Manage.* 10, 32–43.

Hsieh, M., and Tsao, W. (2014). Reducing perceived online shopping risk to enhance loyalty: a website quality perspective. *J. Risk Res.* 17, 241–261. doi: 10.1080/13669877.2013.794152

Inkson, C. (2019). Service failures and recovery in tourism and hospitality: a practical manual. *Tour. Manage.* 70, 176–177. doi: 10.1016/j.tourman.2018.08.010

Jaeger, S. R., Roigard, C. M., David, J., and Vidal, L. (2018). Valence, arousal and sentiment meanings of 33 facial emoji: insights for the use of emoji in consumer research crisis. *Food Res. Int.* 119, 895–907. doi: 10.1016/j.foodres.2018.10.074

Jeon, H. M., Sung, H. J., and Kim, H. Y. (2020). Customers' acceptance intention of self-service technology of restaurant industry: expanding UTAUT with perceived risk and innovativeness. *Serv. Bus.* 14, 533–551. doi: 10.1007/s11628-020-00425-6

Jia, W., and Zhao, Z. (2018). Does service remedy really lead to customer satisfaction: from the perspective of customer emotions. *J. Northeast. Univer.* 20, 44–51.

Jian, Z. Q., and Ke, Y. (2017). Research on the relationship between online shopping service failure, service recovery and customer secondary satisfaction and loyalty. *Manage. Rev.* 29, 175–186.

Khatoon, S., and Rehman, V. (2021). Negative emotions in consumer brand relationship: a review and future research agenda. *Int. J. Cons. Stud.* doi: 10.1111/ijcs.12665

Kim, J., Bojanic, D. C., and Warnick, R. B. (2009). Price bundling and travel product pricing practices used by online channels of distribution. *J. Travel Res.* 47, 403–412. doi: 10.1177/0047287508328658

Kim, W., and Lee, H. (2004). Comparison of web service quality between online travel agencies and online travel suppliers. *J. Travel Tour. Market.* 17, 105–116. doi: 10.1300/J073v17n02_09

Lee, H. H., and Moon, H. (2015). Perceived risk of online apparel mass customization: scale development and validation. *Cloth. Textil. Res. J.* 33, 115–128. doi: 10.1177/0887302X15569345

Lee, S. H., and Jung, K. S. (2015). The strategical approach to reduce negative emotion and complaint behavior in social commerce: focusing on perceived risk. *E Bus. Stud.* 16, 103–122. doi: 10.15719/geba.16.3.201506.103

Li, B. L., and Li, Q. (2007). Analysis of perceived risk and its sources of online consumers. *Econ. Manage.* 78–83.

Li, Y. P. (2014). Research on the impact of switching cost on customer loyalty in the context of service failure. *Technol. Econ. Manage. Res.* 63–66. doi: 10.3969/j.issn.1004-292X.2014.08.014

Liljander, V., and Strandvik, T. (1997). Emotions in service satisfaction. *Int. J. Serv. Indust. Manage.* 8, 148–169. doi: 10.1108/09564239710166272

Ling, L., Dong, Y., Guo, X., and Liang, L. (2015). Availability management of hotel rooms under cooperation with online travel agencies. *Int. J. Hosp. Manage.* 50, 145–152. doi: 10.1016/j.ijhm.2015.07.005

Ling, L., Guo, X., and Yang, C. (2014). Opening the online marketplace: an examination of hotel pricing and travel agency online distribution of rooms. *Tour. Manage.* 45, 234–243. doi: 10.1016/j.tourman.2014.05.003

Liu, F. J., Meng, L., Yang, Q., and Liu, Y. Y. (2019). The effect of matching type of prior remedy and customer participation on service recovery performance from the perspective of responsibility attribution. *Nankai Manage. Rev.* 22, 197–210.

Liu, G. W., and Li, M. G. (2017). Online shopping service remediation quality measurement system and improvement of SERVQUAL model: an experimental analysis perspective based on policy preference. *China Bus. Market* 9, 63–70.

Liu, J., and Zhang, E. (2014). An investigation of factors affecting customer selection of online hotel booking channels. *Int. J. Hosp. Manage.* 39, 71–83. doi: 10.1016/j.ijhm.2014.01.011

Luarn, P., and Lin, H. H. (2003). A customer loyalty model for E-Eervice context. *J. Electro. Comm. Res.* 4, 154–167.

Maxham, J. G. (2001). Service recovery's influence on consumer satisfaction, positive word of mouth, and purchase intention. *J. Bus. Res.* 54, 11–24. doi: 10.1016/S0148-2963(00)00114-4

Mi, L. H., Ryong, K. J., and Han, J. S. (2019). The effect of perceived risk on negative emotions, trust, and investment intentions: focused on investors and potential investors. *Int. J. Tour. Hosp. Res.* 33, 19–31. doi: 10.21298/IJTHR.2019.8.33.8.19

Nikbin, D., Marimuthu, M., Hyun, S. S., and Ismail, I. (2015). Relationships of perceived justice to service recovery, service recovery failure attributions, recovery satisfaction, and loyalty in the context of airline travelers. *Asia Paci. J. Tour. Res.* 20, 239–262. doi: 10.1080/10941665.2014.889028

Nunnally, J. C., and Bemstein, H. T. (1994). *Psyehometrie Theory, 3th Edn.* New York, NY: Me Graw Hill.

Olavarria-Jaraba, A., Cambra-Fierro, J. J., Centeno, E., and Vazquez-Carrasco, R. (2018). Relationship quality as an antecedent of customer relationship proneness: a cross-cultural study between Spain and Mexico. *J. Retail. Consum. Serv.* 42, 78–87. doi: 10.1016/j.jretconser.2018.01.011

Oliver, R. L. (1999). Whence consumer loyalty? *J. Market.* 63, 33–44. doi: 10.1177/00222429990634s105

Ou, Y., and Verhoef, P. (2017). The impact of positive and negative emotions on loyalty intentions and their interactions with customer equity drivers. *J. Bus. Res.* 80, 106–115. doi: 10.1016/j.jbusres.2017.07.011

Pai, C. H., Kuo-Min, K. O., and Santos, T. (2019). A study of the effect of service recovery on customer loyalty based on marketing word of mouth in tourism industry. *Rev. Cercetare Intervent. Soc.* 64, 74–84. doi: 10.33788/rcis.64.6

Pappas, N. (2017). Effect of marketing activities, benefits, risks, confusion due to over-choice, price, quality and consumer trust on online tourism purchasing. *J. Market. Commun.* 23, 195–218. doi: 10.1080/13527266.2015.1061037

Paulhus, D. L., Duncan, J. H., and Yik, M. S. M. (2002). Patterns of shyness in East-Asian and European-heritage students. *J. Res. Pers.* 36, 442–462. doi: 10.1016/S0092-6566(02)00005-3

Peng, J. F., and Jing, F. J. (2005). The impact of service failure attribution on service remediation effect. *Nankai Manage. Rev.* 8, 8–15.

Pinto, I., and Castro, C. (2019). Online travel agencies: factors influencing tourists' purchase decisions. *Tour. Manage. Stud.* 15, 7–20. doi: 10.18089/tms.2019.150201

Poggi, I., and D'Errico, F. (2010). The mental ingredients of bitterness. *J. Multi. User Interfac.* 3, 79–86. doi: 10.1007/s12193-009-0021-9

Ponzi, L. J., Fombrun, C. J., and Gardberg, N. A. (2011). Reptrak pulse: conceptualizing and validating a short-form measure of corporate reputation. *Corpor. Reputat. Rev.* 14, 15–35. doi: 10.1057/crr.2011.5

Preacher, K. J., Rucker, D. D., and Hayes, A. F. (2007). Addressing moderated mediation hypotheses: theory methods, and prescriptions. *Multivar. Behav. Res.* 42, 185–227. doi: 10.1080/00273170701341316

Ren, J. L., Du, H. Y., and Wang, X. F. (2019). The correlation between perceived risk and purchase intention in online retail. *China Circul. Econ.* 33, 63–72.

Rianthong, N., Dumrongsiri, A., and Kohda, Y. (2016). Improving the multidimensional sequencing of hotel rooms on an online travel agency web site. *Electro. Comm. Res. Applicat.* 17, 74–86. doi: 10.1016/j.elerap.2016.03.002

Ribbink, D., Van, R., Allard, C. R., and Liljander, V. (2004). Comfort your online customer: Quality, trust and loyalty on the internet. *Manag. Serv. Qual.* 14, 446–456. doi: 10.1108/09604520410569784

Scridon, M. A., Achim, S. A., Pintea, M. O., and Gavriletea, M. D. (2020). Risk and perceived value: antecedents of customer satisfaction and loyalty in a sustainable business model. *Econ. Res.* 32, 909–924. doi: 10.1080/1331677X.2019.1584043

Sheng, L. (2019). An international review on cooperation relations between online travel agencies and traditional tourism enterprises. *World Reg. Stud.* 28, 202–212.

Sjöberg, L. (2007). Emotions and risk perception. *Risk Manage.* 9, 223–237. doi: 10.1057/palgrave.rm.8250038

Smironva, E., Kiatkawsin, K., Lee, S. K., Kim, J., and Lee, C. H. (2019). Self-selection and non- response bias in customers' hotel ratings: a comparison of online and offline ratings. *Curr. Issues Tour.* 32, 375–389.

Sven, M., Benjamin, Q., and Jan, W. (2015). Don't try harder: using customer inoculation to build resistance against service failures. *J. Acad. Market. Sci.* 43, 512–527. doi: 10.1007/s11747-014-0398-1

Tony, S. M. T. (2013). The marketing role of the Internet in launching a hotel: the case of hotel ICON. *J. Hosp. Market. Manage.* 22, 895–908. doi: 10.1080/19368623.2013.734224

Valenzuela, F. R., and Cooksey, R. (2014). Australian retail banking customers' perceptions of time in a service recovery process. *Contemp. Manage. Res.* 24, 123–146. doi: 10.7903/cmr.11641

Wang, F. H., Gao, L.i., and Pan, Y. (2017). The impact of customer participation on customer satisfaction: the mediating role of perceived risk and the moderating role of self-efficacy. *Finan. Res.* 101–107. doi: 10.3969/j.issn.1000-176X.2017.06.015

Wang, R. (2020). The impact of consumer perceived risk on impulsive online shopping behavior: theoretical and empirical. *Bus. Econ. Res.* 46–50.

Wang, X. H., and Guo, Y. F. (2018). The impact of platform e-commerce reputation on platform seller performance: based on the perspective of customer relationship quality. *J. Southwest Univer. Natl.* 39, 128–135.

Wang, Z. G., Chen, X., and Deng, S. H. (2018). Research on the impact of perceived risk and overconfidence of structured product Investors. *J. Manage. Sci.* 21, 82–93.

Westbrook, R. A., and Oliver, R. L. (1991). The dimensionality of consumption emotion patterns and consumer satisfaction. *J. Consum. Res.* 18, 84–91. doi: 10.1086/209243

Woodside, A. G., Frey, L. L., and Daly, R. T. (1989). Linking service quality, customer satisfaction and behavioral intention. *J. Health Care Market.* 9, 97–108.

Wu, M. L. (2010). *Structural Equation Model–Operation and Application of Amos*. Chongqing: Chongqing University Press, 175–210.

Yao, Y. N., and Deng, C. H. (2017). A study on user satisfaction of online medical and health websites based on perceived risk and service quality model. *China Health Stat.* 34, 331–334.

Yoon, S. M., and Lee, C. K. (2014). Examining the relationships among perceived risk, expectancy-disconfirmation, emotion, satisfaction, and behavioral intention: the case of visit experience in a festival. *J. Tour. Sci.* 38, 153–174.

Yu, K. Z., Li, Y., and Yin, Y. T. (2008). An analysis of the mechanism of the impact of service recovery on customer loyalty. *Finan. Theory Pract.* 29, 104–109.

Zeithaml, V. A., Berry, L. L., and Parasuraman, A. (1996). The behavioral consequences of service quality. *J. Market.* 60, 31–46. doi: 10.1177/002224299606000203

Zhao, Y. S., and Wang, S. H. (2012). The impact of service failure on relationship quality and customer repurchase intention in online shopping: an empirical study based on relationship type regulation. *J. Centr. South Univer.* 18, 124–134.

Zheng, C. H., Zhang, J., Guo, Y. R., Zhang, Y. L., and Qian, L. L. (2019). Disruption and reestablishment of place attachment after large-scale disasters: The role of perceived risk, negative emotions, and coping. *Int. J. Disast. Risk Reduct.* 40:101273. doi: 10.1016/j.ijdrr.2019.101273

Zhong, T. L., Hu, P., and Sun, J. (2011). Discussion on customer behavior intention after service recovery based on external comparison. *Manage. Rev.* 23, 59–67.

Conflict of Interest: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Wei. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Psychology

ORIGINAL RESEARCH
published: 18 June 2021
doi: 10.3389/fpsyg.2021.655612

Chinese App User's Needs Profile: From Questionnaire Measurement to Behavior Analysis

*Zaoyi Sun[1], Liang Xu[2], Qi Zhong[3] and Xiuying Qian[2]**

[1]College of Education, Zhejiang University of Technology, Hangzhou, China, [2]Department of Psychology and Behavioral Sciences, Zhejiang University, Hangzhou, China, [3]College of Mechanical Engineering, Zhejiang University of Technology, Hangzhou, China

With the widespread use of mobile devices, the Apps people install and use could be closely linked to their needs. A precise profile of the needs of the user has become a vital foundation of the experience of the user. Previous studies mainly rely on self-reporting to understand the subjective attitudes of the App user toward a single App. This research combined questionnaire measurement and behavior analysis to profile the needs of the App user from a broader perspective. Based on the theoretical model of previous research studies, study 1 developed a novel needs questionnaire measurement of a Chinese App user, which showed good reliability and validity. In study 2, authorized App usage data were collected to construct the behavioral needs profile of a Chinese user. The results showed that the primary needs of the Chinese user remained a relatively high consistency between the questionnaire and the behavior data. The questionnaire-based and behavioral data-based needs profiles provide a reference for further personalized user experience design.

Keywords: App user, needs, profile, questionnaire, behavior analysis

OPEN ACCESS

Edited by:
Anna Esposito, University of Campania 'Luigi Vanvitelli, Italy

Reviewed by:
Maria Koutsombogera, Trinity College Dublin, Ireland
Riccardo Spolaor, Shandong University, China

***Correspondence:**
Xiuying Qian
xqian@zju.edu.cn

Specialty section:
This article was submitted to Human-Media Interaction, a section of the journal Frontiers in Psychology

Received: *19 January 2021*
Accepted: *13 May 2021*
Published: *18 June 2021*

Citation:
Sun Z, Xu L, Zhong Q and Qian X (2021) Chinese App User's Needs Profile: From Questionnaire Measurement to Behavior Analysis. Front. Psychol. 12:655612. doi: 10.3389/fpsyg.2021.655612

INTRODUCTION

In recent years, the emergence of mobile Internet technology has deeply changed the life style of people. According to a recent report released by the App Annie (2020, Q2) in US, as governments and businesses step up their efforts to stop the coronavirus pandemic, a large number of people started home-based work and learning. As a result, people are spending about 5 h per day on mobile devices on average, an increase of 30% from 2019. The high adoption rate of mobile devices makes Apps an interesting field to study the online communication of people. Based on the uses and gratifications theory (UGT; Katz et al., 1974), the needs of the user were driving forces of social media communication. Recently, some researchers focus on what and how the user chooses Apps to satisfy their needs (e.g., Tanta et al., 2014). For practitioners like App developers, the needs of the user are effective and attractive sources for designing meaningful user experience (Krüger et al., 2017). Thus, how to profile the needs of the user scientifically and effectively has been a topic of widespread concern in academia and industry recently. This study aims to profile the needs of the user from both subjective attitudes and behavior analysis.

Traditional studies usually adopted questionnaires to get quantitative data for the needs of the user. Ji (2013) surveyed Chinese university students about their motivations for using mobile Apps and found that information, social, and entertainment needs were important factors in usage intentions. Kang and Jung (2014) conducted an exploratory factor analysis (EFA) on 398 American and 331 Korean college students. They identified five constructs of the smartphone basic needs scale from the two samples based on Maslow's hierarchy of needs model (Maslow, 1954). Hsiao et al. (2016) found that hedonic needs (HDNs) were the main factor affecting the intentions of the users to use social mobile Apps. The above studies were based on prior theoretical frameworks and focused on specific Apps. However, the collected questionnaire data were limited in sample size and types of participants (Montjoye et al., 2013).

In contrast to quantitative research studies, qualitative research does not rely on prior theories or models and is more flexible in the choice of subjects and data processing. Grounded theory (GT) is one of the most widely used methods in qualitative research studies, enabling the researchers to seek out and conceptualize the latent social patterns and structures from raw data (Glaser and Strauss, 1967). Some studies have adopted this method to explore the psychological needs structure of a person (e.g., Dukic et al., 2015; Miraglia, 2015). One recent research has specifically used GT to study the needs of the App users (Sun et al., 2017). Coding both the interview data from different age groups and the data from App reviews, the researchers yielded eight types of psychological needs, namely, utilitarian, low-cost, security, health, hedonic, social, cognitive, and self-actualization needs (ANs). However, the experience and biases of researchers may influence the results of GT (Woolley et al., 2000). Although the work of Sun et al. (2017) has considered a wide range of user groups and App categories, more quantitative research data can facilitate better validation of the model.

As mentioned above, qualitative and quantitative research methods not only have their own advantages but also have some limitations in the exploration of psychological phenomena (Blondel et al., 2015). Thus, a mixed method has been advocated. Pincus (2004) summarized the application of mixed methods in the field of consumer psychology that is, conducting qualitative research to elicit consumer marketing issues and identifying consumer language for questionnaire development in quantitative research studies. Recently, some research studies have adopted mixed method to study the use of information technologies (e.g., Walsh, 2014) but have rarely applied it in the field of needs of App user. Therefore, in study 1, we developed a needs questionnaire of App user based on the theoretical model generated by GT (Sun et al., 2017). This questionnaire could not only provide quantitative data for the aforementioned needs model but can also be used for the evaluation of the needs of App user.

For user experience design, questionnaires could help researchers understand the subjective attitudes of the user at the stage of needs assessment (Biduski et al., 2020). At the stage of data analysis, the behavior/objective information of the user would be used for constructing the user profile. Most of the previous research studies collected historical information to infer the usage preference of the user directly (Xin et al., 2015). Li et al. (2014) have pointed that the long-term behavioral preference of the user should be stable. As stable behavioral patterns reflect the unique psychological traits of the user (Shoda et al., 1994), these types of features could solve the "black box" problem of big data to some degree and enhance the interpretability of the model. Recently, a few works have aggregated App behavior records to infer the traits of a user (e.g., Nave et al., 2018), but these works focused on personality. To our knowledge, only one previous work explored the stability of the needs-based profiles of an individual with many anonymous App usage records of the users (Sun et al., 2019). However, there were two limitations of this study. First, they only got access to the cellular data of the user and not wireless fidelity (WiFi) data, which led to a deviation from the actual usage behavior. Second, due to the limitations of big data analysis method, this study presented only aggregated results and did not further compare the results with self-report data of the user.

Therefore, in study 2, we recruited volunteers to participate in a user experience improvement project supported by a Chinese handset company. The participants authorized researchers to anonymously access and record their App usage data of the past month from the back-end data process platform of the company. These data included both WiFi and cellular usage records, having higher ecological validity. Then, the data were analyzed based on the measurement proposed by Sun et al. (2019) to profile the needs of the user. For personality, some studies have found the consistency between behavior prediction and self-report results (Kosinski et al., 2013). Therefore, we tested whether the behavioral needs profile of the user was consistent with the questionnaire data. In doing so, the consistency of behavioral and self-report needs profile put insights into the driven roles of the needs of humans in using the App. Also in practice, this research provided both subjective and objective evaluation of the needs profile of the App user, which could be applied to personalized user experience analysis systems in online communication.

STUDY 1: DEVELOPMENT OF NEEDS QUESTIONNAIRE OF APP USER

Methods

Operationalization of the Model

The present study aimed to develop a questionnaire considering needs-related issues of the App user from the aforementioned needs model. The previous research model identified eight types of psychological needs related to App use, namely, utilitarian (e.g., increasing work efficiency and saving time), low-cost (e.g., inexpensive or free), security (e.g., ensuring information security and privacy protection), health (e.g., tracking the physical state and health-related data), hedonic (e.g., fun and leisure experience), social (e.g., facilitating communication and self-expression), cognitive (e.g., information searching and curiosity satisfying), and self-actualization (e.g., improving himself/herself) needs (Sun et al., 2017).

For operationalization of the model, the measurement items were adapted from two aspects. First, after an extensive search for the research studies on the psychological needs, especially in the context of mobile internet, we summarized the relative questionnaires and their constructs. All the measurement items were adapted from prior works of literature. The items of utilitarian needs (UNs) were adapted from the study of Liu et al. (2016). The items of low-cost needs (LCNs) were adapted from the study of Ganesh et al. (2010). The items of security needs (SENs) were adapted from the study of Taormina and Gao (2013). The items of health needs (HENs) were adapted from the study of Zhang et al. (2014). The items of HDNs were adapted from the study of Bondad-Brown et al. (2012). The items of social needs (SNs) were adapted from the study of Haridakis and Hanson (2009). The items of cognitive needs (CNs) were adapted from the study of Wang et al. (2014). The items of self-actualization needs were adapted from the study of Kim et al. (2013). Second, to ensure content validity, we referred to the GT study of Sun et al. (2017) and its detailed online coding appendices. The words of the measurement items of low-cost and self-actualization needs were further adapted from the aforementioned study.

After the above operationalization process, 41 items were measured with a seven-point Likert scale, ranging from 1 (strongly disagree) to 7 (strongly agree). Besides, we adopted the measurement items of the usage attitudes (i.e., "I think smart phone will make life easier in the future"; "I think smart phone will become more and more important in the future") and intentions (i.e., "I think smart phone will be more frequently used in the future.") of the smartphone developed by Kang and Jung (2014). The participants were also asked to provide their demographic information, such as age, gender, and occupation.

The above questionnaire was originally created in English, and the validity of the content was discussed by two user experience researchers at the first step. At the second step, two researchers employed back-translation method of Brislin (1970) to ensure translation equivalence. At the third step, 20 university students that had rich App usage experience were invited to rate the clarity and comprehensibility of the items. As a result, the pool of items was reduced from 41 to 35. The final items are presented in the **Appendix**.

Participants and Procedure

The above questionnaire was released on a Chinese online survey platform, and the respondents that finished the survey were rewarded with ¥ 5. In order to test the factor structure of the research model, the original samples were randomly split into two halves (Sample A and Sample B), using the procedure employed by Asendorpf et al. (2001). EFA would be performed on Sample A (N = 909) and confirmatory factor analysis (CFA) would be performed on Sample B (N = 909).

Table 1 shows the demographic information of the two samples. Compared with the studies recruiting college students, our samples had a wider age and geographical distribution.

TABLE 1 | Description of the samples in study 1.

Items	Measure	Sample A	Sample B
		Number of people (Proportion)	Number of people (Proportion)
Gender	Male	341 (37.5%)	403 (44.3%)
	Female	568 (62.5%)	506 (55.7%)
Age	<18 years	45 (5%)	58 (6.4%)
	18–25 years	723 (79.5%)	633 (69.6%)
	26–40 years	114 (12.6%)	172 (18.9%)
	41–50 years	21 (2.3%)	36 (4.0%)
	>50 years	6 (0.7%)	10 (1.1%)
Top-5 locations (Provinces)	Guangdong	425 (46.7%)	323 (35.5%)
	Zhejiang	238 (26.2%)	230 (25.3%)
	HuNan	32 (3.8%)	40 (4.4%)
	JiangSu	29 (3.2%)	38 (4.2%)
	HeNan	28 (3.1%)	40 (4.4%)

Specifically, the geographical information covered 29 provinces in China, and **Table 1** shows the top five provinces that had the most respondents. In addition, our participants covered different age groups, and almost 70% of them were between 18 and 25. Interestingly, the age distribution was consistent with the age structure of the Chinese Internet users reported by the China Internet Network Information Center (CNNIC, 2018).

Data Analysis and Results

Exploratory Factor Analysis

The questionnaire items in Sample A were preliminary checked for EFA assumptions to be met. The sampling adequacy for performing the analysis was verified through the Kaiser–Meyer–Olkin (KMO) test. The total KMO value was 0.94, which was well above the acceptable limit of 0.60 (Hutcheson and Sofroniou, 1999). Bartlett's test of sphericity ($\chi 2$ = 23925.37, $p < 0.001$) indicated that between-item correlations were sufficient to perform EFA. Then, the data of the items were subjected to principal component analyses (PCAs) and a varimax rotation. In the factor analysis, eight factors were extracted and rotated, consistent with the model of Sun et al. (2017). This eight-factor solution was supported by a graphic scree-test and Kaiser's criterion for number of factors retained (Nunnally and Bernstein, 1994). The eight factors retained accounted for 74.5% of the total variance. **Table 2** shows the results.

Confirmatory Factor Analysis

Confirmatory factor analysis was performed on Sample B (N = 909). We evaluated two competitive models: (1) a unidimensional model with all items loading on a general meaning factor and (2) the hypothesized eight-factor model with correlated factors. The values of normed chi (NC)-square, comparative fit index (CFI), goodness-of-fit index (GFI), normed fit index (NFI), and root mean square error of approximation (RMSEA) were compared with the recommended values. The results show that the eight-factor model indicated a good model fit (see **Table 3**).

Reliability Analysis

Cronbach α coefficient was computed for the internal consistency (IC) analysis. Values greater than 0.80 for the Cronbach α coefficient of the total questionnaire and greater than 0.70 for the Cronbach α coefficient of each factor were acceptable (Nunnally and Bernstein, 1994). The results in **Table 4** show that the estimated reliabilities were acceptable for all eight factors (>0.70). In addition, the Cronbach α coefficient of the total questionnaire was 0.95, which was also above the critical point 0.80 level of reliability. In conclusion, the questionnaire had acceptable homogeneity reliability.

TABLE 2 | Varimax-rotated factor loadings of exploratory factor analysis (EFA; Sample A; *N* = 909).

	SN	HEN	UN	LCN	SEN	HDN	CN	SAN
SN1	**0.76**	0.02	0.19	0.12	−0.05	0.16	0.23	0.11
SN2	**0.81**	0.11	0.07	0.09	0.05	0.17	0.23	0.07
SN3	**0.71**	0.13	0.14	0.10	0.11	0.17	0.16	0.18
HEN1	0.15	**0.76**	0.12	0.06	0.21	0.04	0.09	0.21
HEN2	0.07	**0.89**	0.10	0.06	0.12	0.06	0.08	0.21
HEN3	0.05	**0.89**	0.08	0.06	0.15	0.06	0.05	0.20
HEN4	0.09	**0.82**	0.12	0.10	0.16	0.03	0.09	0.18
HEN5	0.04	**0.70**	0.08	0.10	0.29	0.00	0.03	0.16
UN1	0.11	0.20	**0.81**	0.11	0.17	0.22	0.21	0.11
UN2	0.11	0.19	**0.85**	0.13	0.15	0.27	0.16	0.14
UN3	0.18	0.16	**0.69**	0.17	0.12	0.29	0.22	0.17
UN4	0.16	0.04	**0.70**	0.26	0.07	0.28	0.34	0.16
LCN1	0.16	0.13	0.17	**0.77**	0.02	0.27	0.18	0.13
LCN2	0.12	0.09	0.13	**0.83**	0.10	0.18	0.23	0.15
LCN3	0.05	0.13	0.15	**0.78**	0.15	0.13	0.21	0.07
SEN1	0.04	0.18	0.15	0.15	**0.77**	0.20	0.11	0.19
SEN2	0.05	0.21	0.18	0.16	**0.76**	0.19	0.15	0.22
SEN3	0.04	0.31	0.06	0.03	**0.81**	0.11	0.03	0.24
SEN4	0.03	0.34	0.07	−0.01	**0.77**	0.08	0.04	0.26
HDN1	0.17	0.00	0.13	0.13	0.04	**0.73**	0.23	0.02
HDN2	0.17	0.02	0.16	0.13	0.11	**0.79**	0.22	0.14
HDN3	0.15	0.06	0.22	0.12	0.19	**0.75**	0.23	0.21
HDN4	0.10	0.03	0.13	0.10	0.11	**0.74**	0.28	0.10
HDN5	0.13	0.11	0.21	0.15	0.22	**0.70**	0.17	0.20
HDN6	0.11	0.02	0.23	0.14	0.06	**0.61**	0.41	0.19
CN1	0.15	0.10	0.20	0.20	0.05	0.30	**0.72**	0.16
CN2	0.11	0.08	0.15	0.17	0.13	0.30	**0.71**	0.14
CN3	0.22	0.03	0.16	0.16	0.07	0.26	**0.76**	0.14
CN4	0.16	0.07	0.19	0.15	0.02	0.30	**0.72**	0.16
CN5	0.19	0.13	0.14	0.12	0.12	0.26	**0.67**	0.20
SAN1	0.07	0.20	0.14	0.10	0.15	0.16	0.18	**0.75**
SAN2	0.07	0.21	0.12	0.06	0.19	0.16	0.17	**0.81**
SAN3	0.08	0.25	0.09	0.07	0.20	0.18	0.07	**0.83**
SAN4	0.16	0.19	0.10	0.08	0.16	0.14	0.20	**0.77**
SAN5	0.11	0.25	0.09	0.12	0.24	0.06	0.15	**0.72**

The bold values indicated that each within construct item loading is higher on the measured construct than the cross-loadings on the other items. UN, utilitarian need; LCN, low-cost need; SEN, security need; HEN, health need; HDN, hedonic need; SN, social need; CN, cognitive need; and AN, self-actualization need.

Validity Analysis

From the perspective of content validity, the questionnaire items in study 1 were based on the previous questionnaires of needs and were inspired by the language of the interviews of the App user in a work of GT. The understandability and clarity of the contents were also evaluated by the scholars. In addition, the results of EFA and CFA indicated that the questionnaire had good structural validity. Convergent validity measures were assessed by the average variance extracted (AVE). The results in **Table 4** show that the AVE of all the factors exceeds the threshold of 0.5 (Fornell and Larcker, 1981).

In terms of criterion-related validity, Kang and Jung (2014) found that the more needs of the user were satisfied through the usage of mobile Apps, the stronger was his/her intention to continue to use mobile phones, and the more positive was his/her attitude toward smartphones. Therefore, the participants in study 1 were also measured about their usage intention and attitudes of the smartphone. **Table 5** shows the results of correlation analysis between the needs factors of the mobile App users and their usage intention and attitudes of the smartphone. Each needs dimension was positively correlated with the scores of smartphone use intention and attitude ($p < 0.001$).

STUDY 2: NEEDS PROFILE OF THE USER BASED ON BEHAVIOR ANALYSIS

In study 2, we attempted to construct the needs profile based on the App usage data and to investigate whether the behavioral needs profile of the individual was consistent with the questionnaire data.

Participants and Procedure

In this study, we cooperated with a Chinese handset company for a user experience improvement project. We recruited participants that used the mobile phone of this company in online forums. We informed that the participants that, if they took part in the project, they would allow the researchers to access their App usage records of the past month from the back-end data process platform of the company, including records of the App name and timestamp of each usage. The Apps of the users were all downloaded from the Huawei App market and run on Android systems. These data would enter the analysis stages in the form of anonymity, and the whole analysis stages would be completed only on the platform of the company.

TABLE 3 | Goodness-of-fit indices from confirmatory factor analysis (Sample B; *N* = 909).

Model	χ^2/df (NC)	GFI	CFI	NFI	RMSEA
Unidimensional	15.58	0.54	0.66	0.65	0.13
Eight factors	2.99	0.91	0.96	0.96	0.05
Recommended value	≤3	>0.9			≤0.05

TABLE 4 | Cronbach α coefficient and average variance extracted (AVE) of factors.

	Item numbers	Cronbach α coefficient	AVE
Social needs	3	0.80	0.57
Healthy needs	5	0.92	0.72
Utilitarian needs	4	0.86	0.62
Low-cost needs	3	0.87	0.63
Security needs	4	0.91	0.61
Hedonic needs	6	0.91	0.57
Cognitive needs	5	0.90	0.53
Self-actualization needs	5	0.92	0.80

TABLE 5 | Pearson's correlation between needs factors and usage intention and attitudes of smartphone.

	Usage intention	Attitude toward smartphones
Healthy needs	0.181***	0.218***
Security needs	0.304***	0.304***
Low-cost needs	0.367***	0.386***
Utilitarian needs	0.473***	0.499***
Hedonic needs	0.553***	0.544***
Social needs	0.394***	0.440***
Cognitive needs	0.523***	0.513***
Self-actualization needs	0.333***	0.404***

***Correlation is significant at the 0.001 level (two-tailed).*

The participants that agreed to join in this project needed to complete questionnaire items constructed in study 1 online. At the end of the questionnaire, they were asked to fill in the international mobile equipment identity (IMEI) code of their mobile phone. IMEI is referred as mobile phone "serial number" and is usually used to identify a specific mobile phone in the communication network of an operator (Stutz et al., 2004). All data of the participants were scrutinized and those with the wrong IMEI codes were deleted. Therefore, a valid sample of 30 participants was used for further analysis. Besides, the participant that completed the questionnaire and whose data were successfully collected would be rewarded with ¥50. The above research procedure was reviewed and agreed by the ethics committee of Zhejiang University of Technology. Thirty-two participants who agreed to participate in the study signed an electronic informed consent.

Calculation

Questionnaire Data

The psychological needs questionnaire of the mobile App users in study 1 contained eight measurement items of psychological needs dimensions. Items in each dimension were averaged, with higher scores representing higher levels of the construct.

App Usage Data

First, each participant *n*'s 30-day App usage data were collected from the back-end data process platform of the company, including the name of the App, timestamps, and the number of usage times. Based on the previous research methods (Sun et al., 2017, 2019), the probability distribution of each App k on eight needs dimensions was also obtained as:

$$E^K = \left[e_0^k, e_1^k, \ldots, e_7^k\right]^T$$

where $0 \le e_i^k \le 1, \sum_{7}^{i=0} e_i^k = 1$. The data format after convergence was as follows:

User_id|App id|Needs tag|use_times ($n|k|E^k|C_{nk}$).

Then, we calculated the App *k*'s average use times μ_k and the SD σ_k. After that, for each user *n*, his/her actual use times, *m*, was combined to calculate the *Z* score Z_{nk} of each used App as:

$$Z_{nk} = \frac{m - \mu_k}{\sigma_k}, (n = 1, \ldots 30), (k = 1, 2, \ldots y)$$

where *y* is the number of Apps used by user *k*.

Finally, based on the method of Zhang and Yu (2015), weighted by Z_{nk}, we calculated user *n*'s needs dimension *i*′ s weighted mean score vector P_{ni} as:

$$P_{ni} = \frac{\sum_{k=1}^{y} (Z_{nk} \cdot e_{ik})}{\sum_{k=1}^{y} Z_{nk}}, (n = 1, \ldots 30), (i = 0, 1, \ldots 7), (k = 1, 2, \ldots y)$$

Results

Descriptive Statistics

Table 6 shows the demographic characteristics of the participants. About 86.7% participants were 18–65 years old (age: M = 22.39, SD = 3.08), and 83.4% had more than 3 years experience of using mobile phone. These young and experienced users could bring more information for our research. Besides, the Apps used by these participants contained a wide range of categories, such as social media, games, online-shopping, and so on. Besides, the number of App categories used by the individual ranged from 7 to 13, which contained most categories in Huawei App markets, including social media Apps, game Apps, online shopping Apps, and so on. Therefore, our following findings could be generalized.

Results of Primary Needs of the Individual

As questionnaire and App usage needs scores of the individual had different ranges and calculation rules, instead of a direct comparison of the values, we ranked each type of scores from high to low separately and calculated the primary needs of each participant (the needs ranked first). First, for the questionnaire scores, since more than one needs had the same value, there were in total 34 primary needs among all the 30 participants. As shown in **Table 5**, for all the primary needs of the participants, the top three were utilitarian, low-cost, and HDNs, and the proportion of health and SENs was relatively low. Besides, no participant rated the cognitive needs as their primary needs. Although the self-actualization need was at the highest level in Maslow's hierarchy of needs theory, in this research, this need was rated four times (8%) as the primary need of all participants.

Then, we ranked the scores calculated through the App usage data of the participant. Since, in this case, the eight

TABLE 6 | Demographic information of participants in sample 2 (N = 30).

Measure	Items	Number of people (Proportion)
Gender	Male	9 (30%)
	Female	21(70%)
Age	18–25 years	26(86.7%)
	26–40 years	1(3.3%)
	41–50 years	1(3.3%)
	>50 years	2(6.7%)
Experience of using mobile phone	<12 months	1(3.3%)
	1–3 years	4(13.3%)
	>3 years	25(83.4%)

TABLE 7 | Descriptive statistics for primary needs of the participants (the needs ranked first).

Type of needs	Questionnaire scores		Usage data scores	
	Frequency	Percentage (%)	Frequency	Percentage (%)
Utilitarian	9	26.4%	15	50%
Low-cost	8	23.6%	2	6.7%
Hedonic	7	20.6%	4	13.3%
Social	5	14.7%	3	10%
Self-actualization	3	8.9%	0	0
Cognitive	0	0	3	10%
Security	1	2.9%	0	0
Health	1	2.9%	3	10%

rankings of each participant were all unique, there were in total 30 primary needs among all the participants. As shown in **Table 7**, for all the primary needs of the participants, the top two were utilitarian and HDNs. This result was consistent with Nielsen's report about the primary motivations of the App user (Nielsen and Budiu, 2013). Compared with the results of questionnaire data, we found that although cognitive needs were not regarded as the primary needs in subjective evaluation of any user, it was calculated as the primary needs of three participants (10%) from the App usage data. In contrast, the self-actualization needs were not the primary needs of any participant from the App usage data, although, in subjective evaluation, it was rated as the primary need four times.

Needs Profile Analysis

For each participant, the ranking results of questionnaire measurement and App behavior records were two different data sources to profile his/her needs. Thus, we tried to explore the consistency of needs profiles of an individual between the subjective and objective evaluations. Based on the ranking results above, we listed top-3 needs of each participant based on the questionnaire and the App usage scores, respectively, and calculated the matching degree. As shown in **Figure 1**, the results showed that the matching degree of primary needs reached 83.3%. For the primary and secondary needs, the matching degree fell to 36.7%, while the matching degree of the top three needs was close to the random level. These results suggested that the primary needs of the individual remained relatively highly consistent between subjective questionnaire evaluation and the scores calculated by App usage data. Also, the results suggested that for the more nonmajor needs, the matching degree of the two needs profiles was lower.

Difference Test of Top-3 Needs Score

In order to illustrate the rule of primary needs in the App usage behavior of an individual, we calculated mean values of top-3 needs scores of all the 30 participants, respectively, based on App usage data. **Figure 2** shows the results that the mean value of the primary needs was 6.13, the mean value of the secondary needs was 2.65, and the mean value of the third needs was 1.82. In order to further explain that the proportion of the primary needs among the top-3 needs, we calculated the difference of the mean value of the primary and secondary needs (Value 1) and difference of the mean value of the secondary and third needs (Value 2). T-test was conducted and showed that Value 1 was significantly higher than Value 2, t (29) = 3.25, $p < 0.01$, Cohen d = 0.84. The results are shown in **Figure 3** and suggested that the value of the primary needs had a relatively large proportion of top-3 needs of behavioral needs profile of an individual.

DISCUSSION

This research combined questionnaire measurement and behavior data analysis to profile the psychological needs of the App user. In study 1, based on the previous grounded work and needs theories, we constructed a questionnaire for the needs measurement of the App user and showed good reliability and validity. In study 2, based on App usage data, we calculated the needs scores of the user and compared the consistency between behavioral needs profile and questionnaire data of study 1. The results showed that the primary needs of individual user remained, approximately, highly consistent (86.7%) between subjective evaluation and the behavior data analysis.

In contrast to the general way of questionnaire construction, the theoretical framework of our study came from the previous results of GT. All items could be traced back to the interview text of the user, which enhanced the comprehensibility (Bergman and Hallberg, 2002). Although GT was classical qualitative research that did not necessarily require quantitative verification, some researchers adopted mixed method to support the model of GT with quantitative data (e.g., De Smet et al., 2019). In our research, EFA was performed and eight factors were extracted and rotated, which were consistent with the previous dimensions of the needs model of GT. A CFA further corroborated the eight-factor model compared to the unidimensional solution and verified the theoretical model of Sun et al. (2017) with questionnaire data. Although the core dimensions were consistent with the previous study of GT, the contents of some subdimensions were enriched by literature review. For example, in the previous study, the cognitive needs included "information searching" and "curiosity satisfying" categories. In this research, we enriched the "information searching"

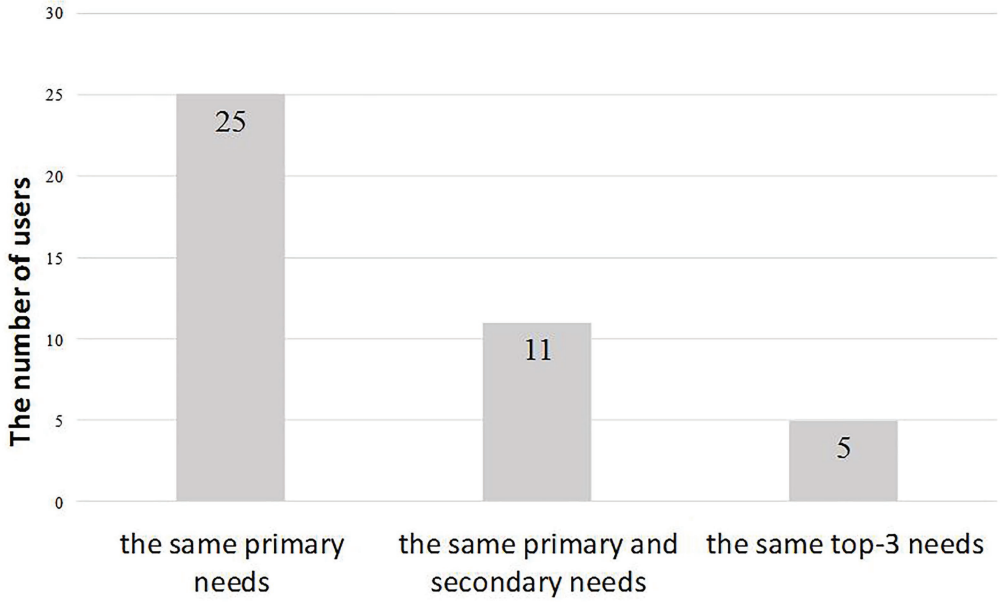


FIGURE 1 | The matching degree of top-3 needs of each participant's.

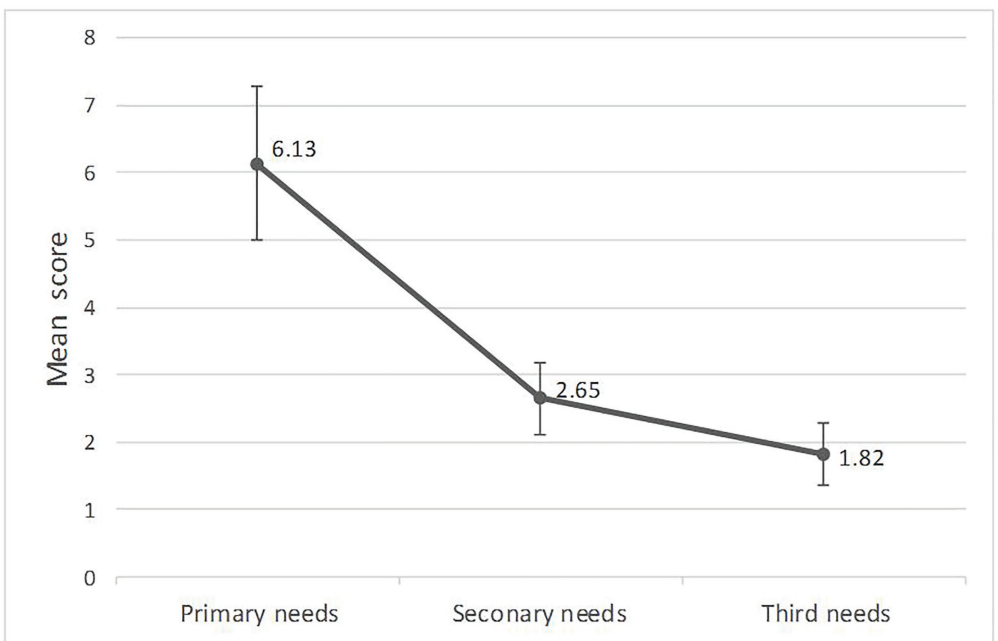


FIGURE 2 | Means of top-3 needs scores of participants based on App usage data.

category and added measurement items based on some related studies (e.g., Gan and Li, 2018).

As for the criterion-related validity analysis results of the questionnaire in study 1, the average scores of the eight needs dimensions were positively correlated with the intention and attitudes of the usage of smartphones ($p < 0.001$). These results could be explained by UGT, that is, the higher the needs of the user was satisfied by mobile Apps usage, the stronger he/she had the intention to use the mobile phone and had the more positive usage attitudes. Specifically, in terms of the correlation coefficient, the correlation coefficients of healthy, security, and self-actualization needs were relatively lower than other needs, which was consistent with the results of GT (Sun et al., 2017). In study 2, scores of these three needs were also relatively lower than other needs in both questionnaire scores and behavior data scores.

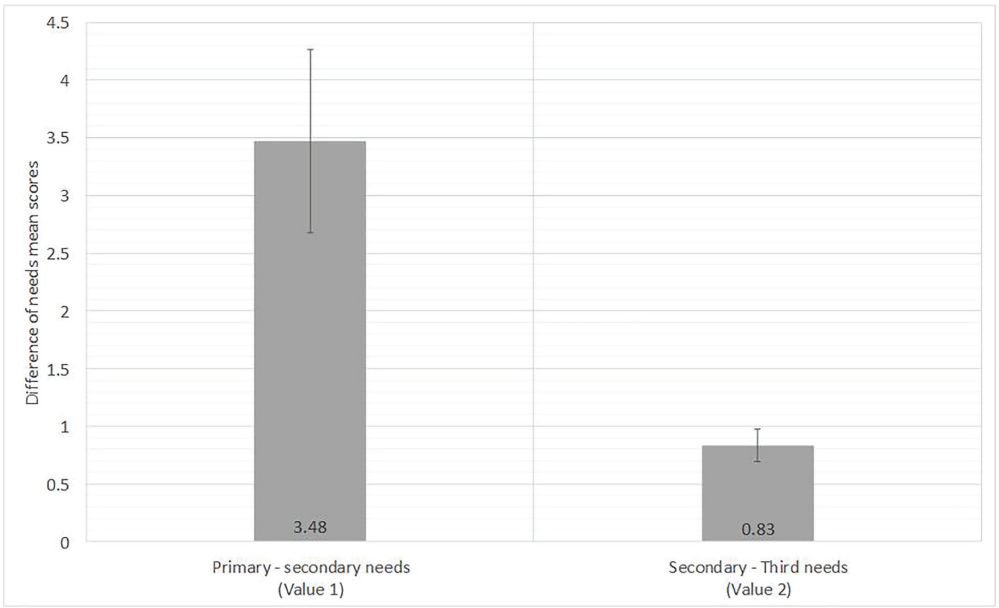


FIGURE 3 | Difference test mean values of top-3 needs of participants.

In addition, for the cognitive needs and self-actualization needs, there were gaps between questionnaire evaluation and behavior data scores in study 2. Pun (2015) mentioned that more and more people used smartphones to obtain information for the satisfaction of their cognitive needs. In the era of fragmented reading, people could easily get a lot of information at their preferred times and places, but few of them would in fact benefit the self-growth of the individual (Liu and Huang, 2016). Therefore, the participants expected to satisfy their self-actualization needs subjectively, but most of the time they just used Apps for information collection with shallow reading. These gaps reflected the expectations of the individual that would influence his/her subjective evaluations. The results would also inspire the App developers to satisfy the self-actualization needs of the user by helping users deal with the challenges associated with fragmented reading.

In study 2, we found that the primary needs of the individual were highly consistent between questionnaire evaluation and App usage data. For the nonmajor needs, the matching degree decreased significantly. Sun et al. (2019) have found that needs-based profiles of App usage of individual were stable across different usage situations. Our research explored the stability of needs profile of the App user across self-report and usage behaviors and identified that the primary needs of the user was traits-like. These results indicated that, as the driving force, the primary needs shaped the App usage preference of the user from attitudes to behaviors. A previous study has proposed that "activated" needs of an individual shaped his/her consumption preference and purchase decision (Griskevicius and Kenrick, 2013). Gender, life experience, developmental stage, culture, and other factors were considered to affect the priority of the individual of activated needs (Kenrick et al., 2010). Our study also found that the value of primary needs of the user had a relatively larger proportion than other needs. Hirsh et al. (2012) also found similar results of personality traits of the individual. They found that, when the ads match more domain traits of the individual, they both rated it more effective and expressed greater purchase intentions. Therefore, as an application of stability, the primary needs of the user could be a core feature in the online recommendation system to profile the user and predict his/her preference.

How to profile the users precisely has been an important issue in the field of user experience (Shalala et al., 2018). Previous studies always profiled that the needs of the users completely depended on behavior record data without subjective report data of the user (e.g., Sun et al., 2019). This data-driven method can rapidly extract patterns in a short time, but it is difficult to interpret the results and to understand how they relate to theory. One way to solve the problem is to import more types of data. For example, research of Maldonado-Mahauad et al. (2018a,b) combined the behavior data collected from an online course learning platform of the users with relevant questionnaire data and identified six different online learning strategies. Our research explored to integrate the bottom-up approach of mining App behavior data with the traditional top-down approach of using validated self-reporting instruments. As a result, we designed two-dimensional needs profiles and combined two types of data sources. As an application, these needs questionnaires and behavior data analysis methods could be chosen by the user experience researchers at App development stage.

Finally, several limitations of this study need to be discussed. First, the sample size in study 2 was relatively small, because many users were cautious about the disclosure of their usage records as the data surely contained personal information and can reflect usage preference. Future studies

should attempt to replicate current findings with more samples to make the results more solid. Second, for the proportions in study 2, few male participants compared to female participants authorized researchers to track their App usage behaviors. Previous studies about self-disclosure have found that females are more likely to share their information than males (e.g., Kays et al., 2012). As the proportions of the sample could not be manipulated, in future studies, we will focus on adopting a more balanced sample for more reliable results. Third, this research focused on App users in China; hence, the results may be limited to the Chinese population. Future studies could explore whether the questionnaires in study 1 and the behavior analysis procedure in study 2 have good effects in different cultures. Finally, this research only adopted two types of data sources to profile the needs of the user. More data sources, such as behavioral experiment data, comment text of the user, and so forth, could be combined to obtain a multidimensional user profile in the future.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Ethics Committee of Zhejiang University of Technology. The patients/participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

AUTHOR CONTRIBUTIONS

ZS developed the study concept and drafted the manuscript. LX performed the testing and data collection. ZS and QZ performed the data analysis and interpretation. XQ provided critical revisions to the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by Digital Intelligent Manufacturing Cooperative Project Funding (KYY-HX-20200166) and Zhejiang University of Technology's Scientific Research Funding (GZ20511080007).

REFERENCES

App Annie (2020). Mobile App Usage Surged 25% YoY in Q3 2020. Available at: https://www.appannie.com/cn/insights/market-data/mobile-app-usage-surged-in-q3-2020/ (Accessed March 21, 2020).

Asendorpf, J. B., Borkenau, P., Ostendorf, F., and Van Aken, M. A. G. (2001). Carving personality description at its joints: confirmation of three replicable personality prototypes for both children and adults. *Eur. J. Personal.* 15, 169–198. doi: 10.1002/per.408

Bergman, B., and Hallberg, L. R.-M. (2002). Women in a male-dominated industry: factor analysis of a women workplace culture questionnaire based on a grounded theory model. *Sex Roles* 46, 311–322. doi: 10.1023/A:1020276529726

Biduski, D., Bellei, E. A., Rodriguez, J. P. M., Zaina, L. A. M., and De Marchi, A. C. B. (2020). Assessing long-term user experience on a mobile health application through an in-app embedded conversation-based questionnaire. *Comput. Hum. Behav.* 104:106169. doi: 10.1016/j.chb.2019.106169

Blondel, V. D., Decuyper, A., and Krings, G. (2015). A survey of results on mobile phone datasets analysis. *EPJ Data Sci.* 4, 1–55. doi: 10.1140/epjds/s13688-015-0046-0

Bondad-Brown, B. A., Rice, R. E., and Pearce, K. E. (2012). Influences on TV viewing and online user-shared video use: demographics, generations, contextual age, media use, motivations, and audience activity. *J. Broadcast. Electron. Media* 56, 471–493. doi: 10.1080/08838151.2012.732139

Brislin, R. W. (1970). Back-translation for cross-cultural research. *J. Cross-Cult. Psychol.* 1, 185–216. doi: 10.1177/135910457000100301

CNNIC (2018). The 42nd Statistical Report on Internet Development in China. China Internet Network Information Center. Available at: http://www.cnnic.net.cn/hlwfzyj/hlwxzbg/hlwtjbg/201808/P020180820630889299840.pdf (Accessed October 31, 2020).

De Smet, M. M., Meganck, R., Van Nieuwenhove, K., Truijens, F. L., and Desmet, M. (2019). No change? A grounded theory analysis of depressed patients' perspectives on non-improvement in psychotherapy. *Front. Psychol.* 10:588. doi: 10.3389/fpsyg.2019.00588

Dukic, Z., Chiu, D. K. W., and Lo, P. (2015). How useful are smartphones for learning? Perceptions and practices of library and information science students from Hong Kong and Japan. *Libr. Hi Tech* 33, 545–561. doi: 10.1108/LHT-02-2015-0015

Fornell, C., and Larcker, D. F. (1981). Structural equation models with unobservable variables and measurement error: algebra and statistics. *J. Mark. Res.* 18, 382–388. doi: 10.1177/002224378101800313

Gan, C., and Li, H. (2018). Understanding the effects of gratifications on the continuance intention to use WeChat in China: a perspective on uses and gratifications. *Comput. Hum. Behav.* 78, 306–315. doi: 10.1016/j.chb.2017.10.003

Ganesh, J., Reynolds, K. E., Luckett, M., and Pomirleanu, N. (2010). Online shopper motivations, and e-store attributes: an examination of online patronage behavior and shopper typologies. *J. Retail.* 86, 106–115. doi: 10.1016/j.jretai.2010.01.003

Glaser, B. G., and Strauss, A. L. (1967). The discovery of grounded theory. *Int J Qual Methods* 5, 1–10. doi: 10.2307/588533

Griskevicius, V., and Kenrick, D. T. (2013). Fundamental motives: how evolutionary needs influence consumer behavior. *J. Consum. Psychol.* 23, 372–386. doi: 10.1016/j.jcps.2013.03.003

Haridakis, P., and Hanson, G. (2009). Social interaction and co-viewing with YouTube: blending mass communication reception and social connection. *J. Broadcast. Electron. Media* 53, 317–335. doi: 10.1080/08838150902908270

Hirsh, J. B., Kang, S. K., and Bodenhausen, G. V. (2012). Personalized persuasion: tailoring persuasive appeals to recipients' personality traits. *Psychol. Sci.* 23, 578–581. doi: 10.1177/0956797611436349

Hsiao, C. H., Chang, J. J., and Tang, K. Y. (2016). Exploring the influential factors in continuance usage of mobile social apps: satisfaction, habit, and customer value perspectives. *Telemat. Inform.* 33, 342–355. doi: 10.1016/j.tele.2015.08.014

Hutcheson, G. D., and Sofroniou, N. (1999). *The Multivariate Social Scientist: Introductory Statistics Using Generalized Linear Models*. New York: Sage Publications, 1–276.

Ji, J. (2013). A study of external factors affecting college students' adoption to use mobile phone in the perspective of TAM. [master's thesis]. Nanjing: University of Nanjing.

Kang, S., and Jung, J. (2014). Mobile communication for human needs: a comparison of smartphone use between the US and Korea. *Comput. Hum. Behav.* 35, 376–387. doi: 10.1016/j.chb.2014.03.024

Katz, E., Blumler, J. G., and Gurevitch, M. (1974). Uses and gratifications research. *Public Opin. Q.* 37, 509–523. doi: 10.1086/268109

Kays, K., Gathercoal, K., and Buhrow, W. (2012). Does survey format influence self-disclosure on sensitive question items? *Comput. Hum. Behav.* 28, 251–256. doi: 10.1016/j.chb.2011.09.007

Kenrick, D. T., Griskevicius, V., Neuberg, S. L., and Schaller, M. (2010). Renovating the pyramid of needs: contemporary extensions built upon ancient foundations. *Perspect. Psychol. Sci.* 5, 292–314. doi: 10.1177/1745691610369469

Kim, J., Lee, S., and Park, Y. (2013). User-centric service map for identifying new service opportunities from potential needs: a case of app store applications. *Creat. Innov. Manag.* 22, 241–264. doi: 10.1111/caim.12037

Kosinski, M., Stillwell, D., and Graepel, T. (2013). Private traits and attributes are predictable from digital records of human behavior. *Proc. Natl. Acad. Sci. U. S. A.* 110, 5802–5805. doi: 10.1073/pnas.1218772110

Krüger, A. E., Kurowski, S., Pollmann, K., Fronemann, N., and Peissner, M. (2017). "Needs profile-sensitising approach for user experience research," in *ACM International Conference Proceeding Series*; November 2017; 41–48.

Li, L., Zheng, L., Yang, F., and Li, T. (2014). Modeling and broadening temporal user interest in personalized news recommendation. *Expert Syst. Appl.* 41, 3168–3177. doi: 10.1016/j.eswa.2013.11.020

Liu, I. L. B., Cheung, C. M. K., and Lee, M. K. O. (2016). User satisfaction with microblogging: information dissemination versus social networking. *J. Assoc. Inf. Sci. Technol.* 67, 56–70. doi: 10.1002/asi.23371

Liu, Z., and Huang, X. (2016). Reading on the move: a study of reading behavior of undergraduate smartphone users in China. *Libr. Inf. Sci. Res.* 38, 235–242. doi: 10.1016/j.lisr.2016.08.007

Maldonado-Mahauad, J., Pérez-Sanagustín, M., Kizilcec, R. F., Morales, N., and Munoz-Gama, J. (2018a). Mining theory-based patterns from big data: identifying self-regulated learning strategies in massive open online courses. *Comput. Hum. Behav.* 80, 179–196. doi: 10.1016/j.chb.2017.11.011

Maldonado-Mahauad, J., Pérez-Sanagustín, M., Moreno-Marcos, P. M., Alario-Hoyos, C., Muñoz-Merino, P. J., and Delgado-Kloos, C. (2018b). "Predicting learners' success in a self-paced MOOC through sequence patterns of self-regulated learning," in *Lifelong Technology-Enhanced Learning*. eds. V. Pammer-Schindler, M. Pérez-Sanagustín, H. Drachsler, R. Elferink and M. Scheffel (Cham: Springer International Publishing), 355–369.

Maslow, A. H. (1954). *Motivation and Personality*. New York: Harper and Row Publishers, 62–27.

Miraglia, D. A. (2015). The experiences of adult women ages 30–44 who log on to facebook daily using their smartphone: A generic qualitative study. Disseration Abstract International Section B, The Science and Engineering. 76, No-Specified. Available at: https://www.proquest.com/docview/1616603967 (Accessed June 4, 2021).

Montjoye, Y.-A., Quoidbach, J., Robic, F., and Pentland, A. S. (2013). "Predicting personality using novel mobile phone-based metrics," in *Social Computing, Behavioral-Cultural Modeling and Prediction. SBP 2013. Lecture Notes in Computer Science. Vol. 7812*. eds. A. M. Greenberg, W. G. Kennedy and N. D. Bos (Berlin, Heidelberg: Springer), 48–55.

Nave, G., Minxha, J., Greenberg, D. M., Kosinski, M., Stillwell, D., and Rentfrow, J. (2018). Musical preferences predict personality: evidence from active listening and facebook likes. *Psychol. Sci.* 29, 1145–1158. doi: 10.1177/0956797618761659

Nielsen, J., and Budiu, R. (2013). *Mobile Usability*. Stuttgart: MITP-Verlags GmbH & Co., KG.

Nunnally, J. C., and Bernstein, I. H. (1994). *Psychometric Theory Third Edition*. *Vol. 1*. New York: McGraw-Hili, Inc., 752.

Pincus, J. (2004). The consequences of unmet needs: the evolving role of motivation in consumer research. *J. Consum. Behav.* 3, 375–387. doi: 10.1002/cb.149

Pun, R. (2015). We chat in the library: promoting a new virtual reference service using a mobile app. *Libr. Hi Tech News* 32, 9–11. doi: 10.1108/LHTN-03-2015-0017

Shalala, R., Dubin, R., Hadar, O., and Dvir, A. (2018). "Video QoE prediction based on user profile," in *2018 International Conference on Computing, Networking and Communications, ICNC 2018*; March 1, 2018; 588–592

Shoda, Y., Mischel, W., and Wright, J. C. (1994). Intraindividual stability in the organization and patterning of behavior: incorporating psychological situations into the idiographic analysis of personality. *J. Pers. Soc. Psychol.* 67, 674–687. doi: 10.1037/0022-3514.67.4.674

Stutz, M., Thomas, V. M., and Saar, S. (2004). "Linking bar codes to recycling information for mobile phones," in *IEEE International Symposium on Electronics and the Environment*; May 10, 2004; 313–316.

Sun, Z., Ji, Z., Zhang, P., Chen, C., Qian, X., Du, X., et al. (2017). Automatic labeling of mobile apps by the type of psychological needs they satisfy. *Telemat. Inform.* 34, 767–778. doi: 10.1016/j.tele.2017.03.001

Sun, Z., Zhang, P., Ji, Z., Chen, C., Wan, Q., and Qian, X. (2019). Cross-situation consistency of mobile App users' psychological needs. *PLoS One* 14:e0215819. doi: 10.1371/journal.pone.0215819

Tanta, I., Mihovilović, M., and Sablić, Z. (2014). Uses and gratification theory—why adolescents use facebook? *Medijska Istraz.* 20, 85–110.

Taormina, R. J., and Gao, J. H. (2013). Maslow and the motivation hierarchy: measuring satisfaction of the needs. *Am. J. Psychol.* 126, 155–177. doi: 10.5406/amerjpsyc.126.2.0155

Walsh, I. (2014). A strategic path to study IT use through users' IT culture and IT needs: A mixed-method grounded theory. *J. Strateg. Inf. Syst.* 23, 146–173. doi: 10.1016/j.jsis.2013.06.001

Wang, D., Xiang, Z., and Fesenmaier, D. R. (2014). Adapting to the mobile world: a model of smartphone use. *Ann. Tour. Res.* 48, 11–26. doi: 10.1016/j.annals.2014.04.008

Woolley, S. R., Butler, M. H., and Wampler, K. S. (2000). Unraveling change in therapy: three different process research methodologies. *Am. J. Fam. Ther.* 28, 311–327. doi: 10.1080/019261800437883

Xin, X., Liu, Z., Lin, C. Y., Huang, H., Wei, X., and Guo, P. (2015). "Cross-domain collaborative filtering with review text," in *IJCAI International Joint Conference on Artificial Intelligence*; July 25, 2015; 1827–1834.

Zhang, M. W. B., Ho, C. S. H., Fang, P., Lu, Y., and Ho, R. C. M. (2014). Usage of social media and smartphone application in assessment of physical and psychological well-being of individuals in times of a major air pollution crisis. *J. Med. Internet Res.* 16:e16. doi: 10.2196/mhealth.2827

Zhang, G., and Yu, E. (2015). "A novel personalized recommendation method in E-business based on kansei image," in *Proceedings of the 2015 International Conference on Automation, Mechanical Control and Computational Engineering.*

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Sun, Xu, Zhong and Qian. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

APPENDIX

Needs Questionnaire Items and Their Sources of Mobile App User

Please evaluate the following 28 questions according to your actual user experience of mobile Apps, from 1 (strongly disagree) to 7 (strongly agree).

Healthy Needs Satisfaction

HEN1: Some mobile Apps enable me to track my physical conditions (Taormina and Gao, 2013; Zhang et al., 2014).
HEN2: Some mobile Apps enable me to exercise the body.
HEN3: Some mobile Apps enable me to enhance physical quality.
HEN4: Some mobile Apps enable me to acquire health guidance.
HEN5: Some mobile Apps enable me to avoid physical fatigue.

Utilitarian Needs Satisfaction

UN1: Using mobile Apps enables me to handle issues easily (Liu et al., 2016).
UN2: Using mobile Apps enables me to handle issues with comprehensive functions.
UN3: Using mobile Apps enables me to find what I want effectively.
UN4: Using mobile Apps enables me to handle issues conveniently.
UN5: Using mobile Apps enables me to deal with tasks quickly (Deleted).

Security Needs Satisfaction

SEN1: Some mobile Apps enable me to feel safe and secure in some situations (Taormina and Gao, 2013).
SEN2: Some mobile Apps enable me to have a sense of trust during use.
SEN3: Some mobile Apps enable me to keep personal information securely.
SEN4: I use some mobile Apps because there will be no privacy concerns.
SEN5: I use some mobile Apps because there will be lower security-related risk (Deleted).

Low-Cost Needs Satisfaction

LCN1: Some mobile Apps enable me to find sales promotion information (Ganesh et al., 2010; Sun et al., 2017).
LCN2: Some mobile Apps enable me to get free products/services.
LCN3: Some mobile Apps enable me to find products/services at reasonable prices.
LCN4: Some mobile Apps enable me to find cheap products/services (Deleted).
LCN5: Some mobile Apps enable me to compare the prices and save money (Deleted).

Social Needs Satisfaction

SN1:Using mobile Apps enables me to belong to a group with the same interests as mine (Haridakis and Hanson, 2009).
SN2: Using mobile Apps enables me to keep in touch with friends in my real life.
SN3: Using mobile Apps enables me to express myself freely.
SN4: Using mobile Apps enables me to communicate with each other (Deleted).
SN5: Using mobile Apps enables me to maintain a daily connection with others (Deleted).

Hedonic Needs Satisfaction

HDN1: I use some mobile Apps when I have nothing better to do (Bondad-Brown et al., 2012).
HDN2: I have fun when using some mobile Apps.
HDN3: The actual process of using some mobile Apps is pleasant.
HDN4: I use some mobile Apps when I am bored.
HDN5: The actual process of using some mobile Apps is relaxed.
HDN6: Some mobile Apps enable me to get sensuous enjoyment.

Cognitive Needs Satisfaction

CN1: Some mobile Apps enable me to access the latest information (Bondad-Brown et al., 2012).
CN2: Some mobile Apps enable me to access the information I wanted to know.
CN3: I use some mobile Apps to provide information.
CN4: I use some mobile Apps to learn about how to do things I have not done before.
CN5: I use some mobile Apps to satisfy my curiosity.

Self-Actualization Needs Satisfaction

SAN1: I use some mobile Apps for career development (Kim et al., 2013).
SAN2: Some mobile Apps enable me to achieve my goals.
SAN3: Some mobile Apps enable me to get my work done.
SAN4: I use some mobile Apps to develop my capacities.
SAN5: I use some mobile Apps for financial and life instructions.

frontiers in Psychology

ORIGINAL RESEARCH
published: 08 July 2021
doi: 10.3389/fpsyg.2021.626040

Cyber-Personality and *Liking* Expression: A Study From WeChat Users in China

*Haojian Li and X.T. Wang**

School of Humanities and Social Science, The Chinese University of Hong Kong, Shenzhen, China

Clicking the *like* button following a post on social media has become a common means of expressing and gathering social support online. Little is known about how *liking* expression is linked and regulated by personality traits and communication motives. Following a preliminary survey ($n = 168$) about the usage of the *like* function on WeChat, a Chinese social media platform, we conducted an online study ($n = 183$) to map the Big-Five personality traits and five communication motives to the frequency (likelihood) of *liking* expression. The results showed that each user had, on average, 385 WeChat friends and spent 2.2 hours and used the *liking* function 1.1 times each day on WeChat. The personality trait *conscientiousness* was negatively related to the *liking* expression ($\beta = -0.505$, $p < 0.05$). In contrast, *agreeableness* promoted the expression of *liking* directly ($\beta = 0.153$, $p < 0.05$) and indirectly via two communication motives, *enjoyment* (a: $\beta = 0.377$, $p < 0.01$; b: $\beta = 0.433$, $p < 0.001$) and *passing time* (c: $\beta = 0.578$, $p < 0.05$; d: $\beta = 0.523$, $p < 0.001$). The *liking* expression may serve as a simple index for understanding dispositional underpinnings of social media networking.

Keywords: cyberpersonality, communication motives, social media, WeChat, *liking* expression, personality traits

OPEN ACCESS

Edited by:
Anna Esposito, University of Campania 'Luigi Vanvitelli, Italy

Reviewed by:
Hui Jing Lu, Hong Kong Polytechnic University, Hong Kong
Dejan Jakimovski, Buffalo Neuroimaging Analysis Center, United States

***Correspondence:**
X.T. Wang
xtwang@cuhk.edu.cn

Specialty section:
This article was submitted to Human-Media Interaction, a section of the journal Frontiers in Psychology

Received: *04 November 2020*
Accepted: *04 June 2021*
Published: *08 July 2021*

Citation:
Li H and Wang X (2021) Cyber-Personality and Liking Expression: A Study From WeChat Users in China. Front. Psychol. 12:626040. doi: 10.3389/fpsyg.2021.626040

INTRODUCTION

Cyber-Personality and Motives of Online Communication

Cyber-personality, human personality in cyberspace, has emerged as a new topic in social and personality psychology. The researchers have suggested that people's online behaviors, including online communication and online self-expression, are regulated by some specific personality traits and motivational factors (e.g., Anolli et al., 2005; Orchard and Fullwood, 2010; Attrill, 2015). In contrast to the vast amount of research on behavioral effects of personality in real life, little is known about how different personality traits manifest in the cyber-environment. The manifestations of personality traits and social motives in the cyber-environment may or may not be consistent with their behavioral effects observed in the actual world. The knowledge about personality effects in actual environments cannot be simply applied to cyber environments. For instance, a timid person in the real world may be highly aggressive in the cyber world. As the first step of the investigation, we intend to focus on a common and measurable online behavior and map it to specific personality traits as traditionally measured by the Big-Five personality inventory. In addition, we also map the online behavior to motivation constructs that drive and regulate social networking.

The online behavior examined in this research is the action of clicking the *like* button following a WeChat post, including, visual, auditory, or text message. WeChat is the most popular social media platform in China with over 1.2 billion monthly active users worldwide as of the first quarter of 2020, according to Statista.com. WeChat plays a dominant role in China's social media and influences people's online behavior, it enhances trust and personal

contacts while transforming opportunities for larger online group formation (Harwit, 2017). WeChat users join friend circles, called "moments" which is similar to a Twitter post. The *like* function is one of the most common tools of expressing and gathering social support on WeChat. For this study, we chose this behavior because it is used readily by all the WeChat users to express social support. Secondly, the *like* clicking behavior is equally tractable and quantifiable for all its users.

In this research, we first establish a behavioral profile of *liking* behavior followed by a second study to analyze the relationship between the *liking* behavior and personality (i.e., the Big-Five personality traits) and motives underlying social networking (i.e., enjoyment seeking, pleasing others, and interpersonal relationship formation). Thus, the *like* expression may be used as a fast and frugal index of unique personality and social motive profiles. In addition, we examined which factors, *dispositional* or *situational*, are more significant regulators of the *like* expression behavior in social media.

The *Like* Function in Communication on Social Media

The *like* function of WeChat provides a ready means for the users to selectively show individualized social support at their fingertips. Compared to *comment* on a post, which is a cognitively triggered behavior, *like* expression is more affectively driven (Kim and Yang, 2017). In general, the use of the *like* button reflects users' enjoyment of, agreement with, or interest in a specific post (Basalingappa et al., 2016). People within a social media group are not only *liking* others but also care about receiving *likes* from others. For the senders, *likes* signal social acceptance and maintain interpersonal relationships. For the receivers, *the likes* catch their attention with low cognitive load and have become a new source of social reward (Scissors et al., 2016). For adolescents, *likes* are considered a sign of social approval (Martinez-Pecino and Garcia-Gavilán, 2019). In particular, adolescents with little social support show a significant preference for these kinds of online social interactions (Leung, 2011). In general, *likes* serve as a cue of status and popularity (Blease, 2015). However, the overuse of social media may result in distress by increasing communication overload and reducing self-esteem (Chen and Lee, 2013).

Previous studies have primarily focused on the joint use of *liking*, *commenting*, and *sharing* and their communicational functions, based on data from Facebook, Twitter, or Instagram (Guo and Sun, 2020; Luo et al., 2020; Obamiro et al., 2020). The present study focuses on the link between the *like* expressions on WeChat and users' basic personality traits and communication motives. First, in a preliminary survey, we glean the behavioral profile of WeChat use in general and *liking* expression in particular. We then map personality traits and communication motives to the frequency of *like* expression.

Personality Traits and Communication Motives and *Liking* Expression

To map the *like* behavior onto basic personality traits and communication motives of the WeChat users, we chose the widely-accepted Big-Five personality traits (Goldberg, 1992; John and Srivastava, 1999) as our personality measures. In user-generated online use of social media, users' personality traits play a crucial role in their engagement in and selection of social media.

Extraversion and openness to experience have been found to increase online activities and engagement in social media (Ross et al., 2009; Correa et al., 2010). However, other studies found that some extroverts may use social media for social enhancement and some introverts may use social media to compensate for the lack of social interactions in face-to-face contexts (Zywica and Danowski, 2008). For neuroticism, some researchers believe that a high level of neuroticism had a greater impact on messaging use and online texting. Neurotic individuals reported stronger addictive tendencies in mobile phone use (Ehrenberg et al., 2008). However, other studies showed that neuroticism was unrelated to online activities and engagement in social media (Ross et al., 2009). For agreeableness, it was reported that a low score was associated with low involvement in social media (Landers and Lounsbury, 2006) and a higher intention for maintaining existing relationships (Horzum, 2016).

Regarding motivations of the consumers of social media, the literature suggests two major motives, enjoyment seeking and interpersonal relationship maintenance, derived from an influential theory in the field of communication, *uses and gratifications theory* (Rubin, 1994; Ruggiero, 2000; Quan-Haase and Young, 2010; Whiting and Williams, 2013). The theory explains how people use social media to gratify their needs. According to the theory, users select different media based on their motivational needs. In a more recent study, Lee et al. (2016) identified from the literature five types of motives that make the users click the *like* button on Facebook. Of the five communication motives (i.e., enjoyment, pleasing others, monetary incentive, pass time, and interpersonal relationship), enjoyment, and interpersonal relationship are the two most salient motives. A cross-cultural study with Chinese users on WeChat (Gan and Wang, 2015) also confirmed that the two major types of gratifications, besides finding content information, were hedonic and social gratifications. In the present study, we focused on the aforementioned five communication motives because they are unique for social media use and also overlap with basic human needs (Maslow, 1954).

Quan-Haase and Young (2010) reported that Facebook use is associated more with having fun and knowing about the social activities occurring in one's social network, whereas *instant messaging* is geared more toward relationship maintenance and development. For this study, we examine not only what type of need the *like* function on WeChat fulfill but also how motives and personality traits separately or jointly affect the users to click the *like* button on WeChat.

Our research questions derived from the aforementioned discussion are as follows:

RQ1. How do people use the *like* function on WeChat?
RQ2. Is there any particular personality trait that regulates the use of *like* expression?
RQ3. What communication motives are driving the use of the *like* function?

Although we do not have strong predictions regarding the aforementioned research questions, we hypothesize that the use of the *like* function on WeChat will be affected by the user's personality and communication movies. In particular, two of the Big-Five personality traits, agreeableness and conscientiousness would likely affect the use of *liking* expressions. While agreeableness may increase the use of the *like* function (e.g., Landers and Lounsbury, 2006), conscientiousness may reduce the use of the *like* function due to a more cautious and responsible attitude. In addition, we predict that the effects of agreeableness on the use of the *like* function would be mediated by communication motives, particularly, the enjoyment-seeking motive, as suggested by some previous findings (Lee et al., 2016).

MATERIALS AND METHODS

Participants

In this research, two samples were collected from university students in Beijing. We conducted a preliminary survey to learn how the young participants use WeChat and the *like* function ($n = 168$, 50 male participants). For the main study, the sample consisted of 190 volunteer participants who use WeChat daily. Seven participants failed to complete the survey and were excluded from data analysis. The age of the main sample ($n = 183$) ranged from 17 to 23 years, with an average age of 18.3 $\pm$ 0.91 years. This research was endorsed by the IRB of the university and conducted before the outbreak of the COVID-19 pandemic.

Measures

Preliminary Survey Questionnaire

The questions in the survey gathered basic information about the usage of WeChat and the *like* function, including the number of friends on WeChat, the time spent each day on WeChat, the frequency of using the *like* function, the percentages of the WeChat posts that the participants agree, trust, and considered useful, etc. (for more details see **Appendix**).

Big-Five Personality Scale

In the main study, we measured the Big-Five personality traits using the 60 items of the NEO-FFI inventory (Costa and McCrae, 1992). The inventory assesses five personality traits: extraversion, agreeableness, conscientiousness, emotional stability, and openness to experience. The Chinese version of this inventory had been used and tested previously (Yao and Liang, 2010). Participants responded to the items on a 5-point Likert scale (1 = strongly disagree to 7 = strongly agree). The Cronbach's alpha coefficients in the present study were 0.86 for extroversion, 0.83 for conscientiousness, 0.70 for agreeableness, 0.66 for openness, and 0.81 for neuroticism (emotional stability).

Communication Motives for Using the *Like* Function

We adopted a 21-item inventory developed by Lee et al. (2016) to measure how the *like* expression on WeChat is related to the following five communication motives: enjoyment, pleasing others, monetary incentive, passing time, and interpersonal relationship. The questionnaire begins with an incomplete statement "I click *like* on WeChat because …" followed with 21 items, such as "I enjoy the content," "It helps me fit in with a group of people," "It helps me receive a bargain deal," "It helps me pass time when I am bored.", and "It helps me improve relations with friends." The participants rated these statements on a 7-point scale with 1 meaning strongly disagree and 7 meaning strongly agree. In this scale, the Cronbach's α is 0.79 for *enjoyment*, 0.87 for *pleasing others*, 0.93 for *money incentive*, 0.82 for passing time, 0.91 for *interpersonal relationship*.

Dependent Variable

The dependent variable was a frequency rating of *like* expression on WeChat. The participants were asked: "How frequently do you use the *like* function on WeChat?" The response was measured on a Likert-type scale with the following response anchors: 1 = never, 2 = rarely, 3 = occasionally, 4 = sometimes, 5 = often, 6 = usually, 7 = always (Vagias, 2006). We chose this Liker-type response instead of an absolute estimation such as the number of *like* clicking each day which is likely to vary from time to time and be subjective to memory biases. The current measure is a numerical likelihood estimate in comparison to other WeChat users.

Procedure and Data Analysis

We first conducted a preliminary survey to obtain a WeChat user profile. In the following main study, the questionnaires were distributed online. The consent for participation was obtained before answering the questions. The personality and motivation measures were presented in a balanced order.

Descriptive statistics, Pearson's correlations with Benjamini-Hochberg correction, were calculated. We then conducted a hierarchical regression analysis to reveal the unique contributions of each set of factors (i.e., Big-Five personalities, communication motives) to the model, followed by a path-analysis to sort out the relationships between Big-Five personality traits, communication motives, and *liking* expression on WeChat.

RESULTS

The Results From the Preliminary Survey

To answer RQ1, the preliminary survey gathered basic information to obtain a WeChat user profile. We also asked the participants for their basic attitude toward, perception of, and emotional experience with the *like* function on WeChat (see **Appendix**).

In general, the frequency of using *like* is much higher than using some other functions (e.g., retweet, comment, and create a new post). The frequency of *liking* behavior was positively correlated with other endorsement behaviors such as retweets. The use of the *like* function was considered a means of expressing immediate positive feedback.

Each user had, on average, 385.4 $\pm$ 322.5 friends on WeChat and spent 2.21 $\pm$ 1.19 hrs and used the *liking* function 1.1 $\pm$ 0.52 times each day on WeChat. Of the posts consumed by the participants, 62% were considered trustworthy, 57% were agreeable, and 44% were considered meaningful. The participants

TABLE 1 | Means, standard deviations, intercorrelations between sex, age, personality traits, communication motives, and liking expression, and *P*-values with Benjamini–Hochberg correction.

		1	2	3	4	5	6	7	8	9	10	11	12	13
1	Gender	1												
2	Age	0.04	1											
3	Openess	−0.08	−0.05	1										
4	Agreeableness	0.05	−0.06	−0.02	1									
5	Conscientiousness	−0.04	−0.10	0.02	0.30**	1								
6	Extroversion	0.03	−0.15**	−0.05	0.46**	0.47**	1							
7	Neuroticism	0.07	0.03	0.03	−0.41**	−0.38**	−0.56**	1						
8	Enjoyment	−0.19**	−0.08	0.00	0.20**	0.05	0.13	−0.01	1					
9	Pleasing others	0.03	−0.06	−0.09	0.16*	−0.15*	0.06	−0.10	0.31**	1				
10	Money incentive	0.06	−0.05	−0.19**	0.13	−0.02	0.02	−0.10	0.19**	0.49**	1			
11	Passing time	−0.11	0.01	−0.15*	0.16*	−0.1	0.08	0.02	0.15*	0.38**	0.21**	1		
12	Interpersonal Relation	−0.11	−0.10	−0.08	0.18*	−0.07	0.001	0.02	0.49**	0.64**	0.30**	0.38**	1	
13	*Liking* Expression	−0.14	−0.02	−0.08	0.21**	−0.15*	0.06	0.02	0.29**	0.34**	0.16*	0.54**	0.35**	1
	p-values§	ns	ns	ns	0.015	ns	ns	ns	0.00036	0.00001	ns	0.00001	0.00001	
Mean		1.72	18.84	3.38	3.397	3.43	3.3	2.98	5.30	3.25	3.18	3.57	4.43	4.02
SD		0.45	0.91	0.42	0.444	0.57	0.64	0.61	0.83	1.23	1.75	1.66	1.33	1.84

N = 183.
**p < 0.05; ** p < 0.01.*
§p-values with Benjamini–Hochberg correction.
ns, not significant.

would use the *like* function only when they liked or strongly liked a post. The frequency rating exceeded the middle point on the response scale (see **Appendix**).

These results suggest that WeChat users did not abuse the *like* function which is easy to use at their fingertips. The rate of re-posting a message was even lower, which was about once every 3 days.

Regression Analysis of Personality Traits, Motives, and the *Like* Expressions

Table 1 shows the means, standard deviations, and inter-correlations between the Big-Five personality traits, five communication motives, and the frequency rating of *like* expression on WeChat. To control the family-wide type-1 error, we conducted Benjamini-Hochberg correction to the predictors of *Like* expression (see *p*-values with Benjamini–Hochberg correction in **Table 1**). The frequency rating of *like* expression was positively correlated with the personality trait *agreeableness* and all five communication motives. and negatively correlated with the personality trait *conscientiousness* (not significant after Benjamini–Hochberg correction).

To test our hypothesis regarding RQ2 and RQ3 and to further sort out the relative contributions of the personality and motivational factors on *like* expression, we conducted hierarchical regression analysis.

Table 2 shows the results of hierarchical regression analysis with the frequency rating of *liking* expression as the dependent variable. We conducted a hierarchical regression analysis to reveal the unique contributions of each set of factors (i.e., Big Five personalities, communication motives) to the model. We chose hierarchical regression to allow hypothesis-based placement of variables into the model instead of automated and algorithm-based analysis, such as stepwise regression. Gender and age were entered in the first step, followed by adding the personality traits at the second step, and the five communication motives in the third step. Communication motives were placed into the hierarchical regression following the personality measures based on the consideration that the motives of communication are more domain-specific than the Big-Five personality traits and are expected to mediate the effects of personalities.

Concerning the *liking* expression, age and gender only accounted for 2% of the variance in the dependent variable. No sex-personality trait interaction or sex-communication motive interaction effects were detected.

The predictability of the model improved significantly by adding the personality traits, which accounted for a further 14.2% of the variance. The motivation traits accounted for another 37.8% of the variance. The total percentage of the variance accounted for by the full model reached 54%. In **Table 2**, we reported R^2 and Adjusted R^2 for model fitting and tested multicollinearity using the variance inflation factor (VIF). The VIF scores (far right column) were all lower than 2, indicating that the predictors included in the multiple regression were not highly correlated with each other.

Direct and Indirect Effects of Personality Traits and Communication Motives on the *Liking* Expression

To further test our hypothesis about mediation effects of motivation between personality and the use of the *like* function, we conducted path analyses using M-plus and revealed

TABLE 2 | Hierarchical regression results of the contributions of age and gender (first step), personality traits (second step), and communication motives (third step) to *liking* expression on WeChat.

Model		Unstandardized coefficients	Standardized coefficients	VIF
		β	*β*	
Liking expression on WeChat				
1	Gender	−0.566	−0.139	
	Age	−0.026	−0.013	
2	Gender	−0.743	−0.183*	1.028
	Age	−0.023	−0.011	
	Neuroticism	0.359	0.12	
	Extroversion	0.3	0.105	
	Conscientiousness	−0.806	−0.250**	1.337
	Agreeableness	1.231	0.297***	1.341
	Openness	−0.384	−0.087	
3	Gender	−0.331	−0.081	
	Age	−0.011	−0.006	
	Neuroticism	0.133	0.045	
	Extroversion	0.041	0.014	
	Conscientiousness	−0.456	−0.142	
	Agreeableness	0.669	0.162*	1.454
	Openness	−0.078	−0.018	
	Pleasing others	0.156	0.104	
	Money incentive	−0.033	−0.032	
	Passing time	0.477	0.431***	1.319
	Interpersonal relation	0.001	0	
	Enjoyment	0.341	0.153*	1.456
	R^2 step 1	0.02	Adjusted R^2 step 1	0.09
	R^2 step 2	0.142	Adjusted R^2 step 2	0.107
	R^2 step 3	0.378	Adjusted R^2 step 3	0.334
	R^2 total	0.54	Adjusted R^2 total	0.531

N = 183.
**p < 0.05; **p < 0.01; ***p < 0.001.*

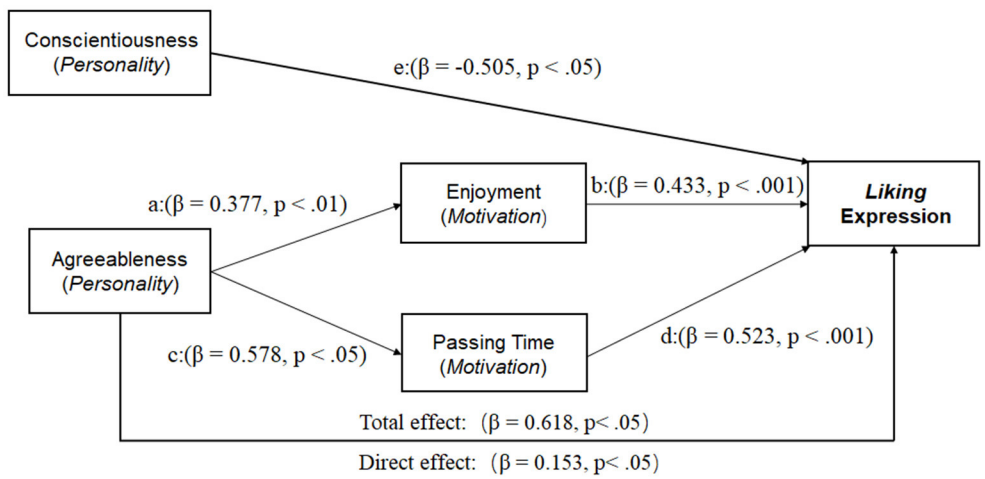


FIGURE 1 | Relationship between personality, motivation, and *liking* expression.

direct and indirect effects of personality and motivation on *liking* expressions.

As shown in **Figure 1**, the personality trait *conscientiousness* was negatively related to the *liking* expression ($\beta = -0.505$, $p < 0.05$), suggesting that higher conscientiousness makes a public expression of social support more serious and less casual. In contrast, another Big-Five personality trait *agreeableness* promoted the expression of *liking* directly (e: $\beta = 0.153$, $p < 0.05$) and indirectly via two communication motives, enjoyment, and passing time. First, the participants who had higher conscientiousness perceived themselves to be less likely to use the liking expression. This direct effect suggests that agreeableness lowers the threshold for a person to express social support in a cyber-environment.

Second, the positive effects of agreeableness on the *liking* expression were mediated by two communication motives. One of the mediating motives was the enjoyment motive, with a path from agreeableness to enjoyment motive (a: $\beta = 0.377$, $p < 0.01$) and then from the enjoyment motive to the *liking* expression (b: $\beta = 0.433$, $p < 0.001$). A bias-corrected bootstraps confidence interval for the indirect effect (ab = 0.163) based on 5,000 bootstrap samples was entirely above zero (0.012–0.315). This mediation effect suggests that using the *liking* function on social media was a pleasant experience and self-rewarding.

The second mediating factor was the motive of passing time, with a link from agreeableness to the motive (c: $\beta = 0.578$, $p < 0.05$) and a final link from the motive to the *liking* expression (d: $\beta = 0.523$, $p < 0.001$). This indirect pathway shows that the *liking* expression on social media provides a way of killing time. A bias-corrected bootstraps confidence interval for the indirect effect (cd = 0.302) based on 5,000 bootstrap samples was entirely above zero (0.012–0.592).

The effect of agreeableness on like expression was reduced after controlling the two motives but remained statically significant (see the outer path in **Figure 1**, $p < 0.05$). Therefore, the result indicates a partial mediation effect between agreeableness and the *liking* expression via the two communication motives. Moreover, the mediating effects of the two motives were especially stronger for a personality of higher agreeableness and lower conscientiousness. The whole model with two partial mediating pathways and one direct pathway had a significant fit of data ($R^2 = 0.337$, $p < 0.001$).

GENERAL DISCUSSION

We derived from the data a four-factor model of dispositional effects on the *lik*ing expression in communication on social media (see **Figure 1**). Consistent with our hypothesis, the personality traits *agreeableness* and *conscientiousness* had, respectively, positive and negative effects on *like* expression. Moreover, the hypothesis that the communication motives mediate the effects of personality traits on the *liking* expression is partially supported where only a single trait (agreeableness) was mediated by leisure-related communication motives.

What the *liking* function is not about should be also important to learn. *Liking* expressions are neither associated with introversion-extraversion nor emotional stability nor openness to experience. In the present study, the sores of extraversion were not significantly correlated with the use of the *like* function on WeChat. The subjective frequency estimate of *like* expression did not distinguish introverts from extroverts. Both introverts and extroverts may benefit from using the *like* function for different reasons. Introverts may benefit from the *like* function to build up their social relations without face-to-face interactions while extroverts may use the *liking* expression for social enhancement (Zywica and Danowski, 2008).

The use of the *liking* function on social media is neither for simply pleasing others nor for utilitarian purposes. In a follow-up study, we asked a group of student participants about the time they spent and the number of the *like* expressions they used each day on WeChat during the COVID-19 pandemic. The results showed that both measures were not statistically different before and after this environmental change. Thus, the use of the *liking* function on social media was not easily susceptible to situational changes in one's environment, indicating it is more of an index of dispositional than situational factors.

Overall, our results indicate that the *liking expression* is more dispositional than situational. It is regulated by personality traits and communication motives rather than situational factors (e.g., environmental changes, health risks, and presumably emotional state due to the COVID-19 pandemic). The findings of the present study suggest that people did not abuse the easy function of *liking* expression on social media. Instead, expressing such social support is regulated by conscientiousness and agreeableness, and mediated by communication motives of passing time and enjoyment seeking.

Limitations and Future Directions

This research has its limitations in its design. First, a clear limitation of this study is its limited demographic sample with little age variation. Our sample was homogeneous, including mainly young university students. These participants are likely to be more active in cyberspace, whose communication motives may be different from other user groups. Second, our measure of the use of the *like* function was a likelihood estimate in comparison to other users in one's friend circle.

People use the *like* function on WeChat for both light-hearted topics and more serious topics. It is plausible and can be examined in future studies that agreeableness is associated more with light-hearted topics while conscientiousness is associated more with serious topics.

Future studies should adopt multiple measures, including the actual frequency measure, attitude toward using the *like* function, and self-confidence measures of social media networking. A comprehensive measure of *liking* expressions may yield a more reliable and valid index for the cyber-personality profiling of the users of social media.

CONCLUSION

The current study is the first to build the user profile of liking expressions on WeChat and to link the *liking* expression to its underlying personality traits and communication motives. The results showed that the *liking* expression is only linked with limited personality traits (conscientiousness and agreeableness) and communication motives (passing time and enjoyment seeking) and is more sensitive to these dispositional factors than situational factors, such as changes due to the COVID-19 pandemic. The unique user profile associated with the *liking* expression makes it a simple and readily available tool for predicting particular communication needs and personal characteristics of social media consumers. The frequency of *liking* expressions reveals the dispositional underpinnings of individual users and social media groups.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethics Committee of the School of Humanities and Social Science, Chinese University of Hong Kong, Shenzhen. The participants provided their informed consent to participate in the study.

AUTHOR CONTRIBUTIONS

HL: methodology, investigation, data curation and analysis, and writing- original draft preparation. XW: conceptualization, methodology, validation, resources, data analysis, writing-reviewing and editing, project supervision, and funding acquisition. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported, in part, by the National Natural Science Foundation of China under Grant Number NSFC 31971025 to XW.

REFERENCES

Anolli, L., Villani, D., and Riva, G. (2005). Personality of people using chat: an on-line research. *Cyberpsychol. Behav.* 8, 89–95. doi: 10.1089/cpb.2005.8.89

Attrill, A. (2015). *Cyberpsychology*. Oxford: Oxford University Press.

Basalingappa, A., Subhas, M. S., and Tapariya, R. (2016). Understanding likes on Facebook: an exploratory study. *Online J. Commun. Media Technol.* 6:234. doi: 10.29333/ojcmt/2566

Blease, C. R. (2015). Too many 'friends,' too few 'likes'? Evolutionary psychology and 'Facebook depression.' *Rev. Gen. Psychol.* 19, 1–13. doi: 10.1037/gpr0000030

Chen, W., and Lee, K.-H. (2013). Sharing, liking, commenting, and distressed? The pathway between Facebook interaction and psychological distress. *Cyberpsychol. Behav. Soc. Netw.* 16, 728–734. doi: 10.1089/cyber.2012.0272

Correa, T., Hinsley, A. W., and Zúñiga, H. G. D. (2010). Who interacts on the Web?: the intersection of users' personality and social media use. *Comput. Human Behav.* 26, 247–253. doi: 10.1016/j.chb.2009.09.003

Costa, P. T., and McCrae, R. R. (1992). *Revised NEO Personality Inventory (NEO-PIR) and NEO Five Factor Inventory (NEO-FFI) Professional Manual*. Odessa, FL: Psychological Assessment Resources.

Ehrenberg, A., Juckes, S., White, K. M., and Walsh, S. P. (2008). Personality and self-esteem as predictors of young people's technology use. *CyberPsychol. Behav.* 11, 739–741. doi: 10.1089/cpb.2008.0030

Gan, C., and Wang, W. (2015). Uses and gratifications of social media: a comparison of microblog and WeChat. *J. Syst. Inform. Technol.* 17, 351–363. doi: 10.1108/JSIT-06-2015-0052

Goldberg, L. R. (1992). The development of markers for the Big-Five factor structure. *Psychol. Assess.* 4, 26–42. doi: 10.1037/1040-3590.4.1.26

Guo, M., and Sun, F. S. (2020). Like, comment, or share? Exploring the effects of local television news Facebook posts on news engagement. *J. Broadcast. Electron. Media* 64, 736–755. doi: 10.1080/08838151.2020.1851125

Harwit, E. (2017). WeChat: social and political development of China's dominant messaging app. *Chin. J. Commun.* 10, 312–327. doi: 10.1080/17544750.2016.1213757

Horzum, M. B. (2016). Examining the relationship to gender and personality on the purpose of Facebook usage of Turkish university students. *Comput. Human Behav.* 64, 319–328. doi: 10.1016/j.chb.2016.06.010

John, O. P., and Srivastava, S. (1999). The Big Five trait taxonomy: history, measurement, and theoretical perspectives. *Handbook Person. Theory Res.* 2, 102–138.

Kim, C., and Yang, S. U. (2017). Like, comment, and share on Facebook: how each behavior differs from the other. *Public Relat. Rev.* 43, 441–449. doi: 10.1016/j.pubrev.2017.02.006

Landers, R. N., and Lounsbury, J. W. (2006). An investigation of Big Five and narrow personality traits in relation to Internet usage. *Comput. Human Behav.* 22, 283–293. doi: 10.1016/j.chb.2004.06.001

Lee, S.-Y., Hansen, S. S., and Lee, J. K. (2016). What makes us click "like" on Facebook? Examining psychological, technological, and motivational factors on virtual endorsement. *Comput. Commun.* 73, 332–341. doi: 10.1016/j.comcom.2015.08.002

Leung, L. (2011). Loneliness, social support, and preference for online social interaction: the mediating effects of identity experimentation online among children and adolescents, *Chin. J. Commun.* 4, 381–399. doi: 10.1080/17544750.2011.616285

Luo, T., Freeman, C., and Stefaniak, J. (2020). "Like, comment, and share"—professional development through social media in higher education: a systematic review. *Educ. Technol. Res. Dev.* 68, 1659–1683. doi: 10.1007/s11423-020-09790-5

Martinez-Pecino, R., and Garcia-Gavilán, M. (2019). Likes and problematic Instagram use: the moderating role of self-esteem. *Cyberpsychol. Behav. Soc. Netw.* 22, 412–416. doi: 10.1089/cyber.2018.0701

Maslow, A. H. (1954). *Motivation and Personality*. New York, NY: Harper

Obamiro, K., West, S., and Lee, S. (2020). Like, comment, tag, share: Facebook interactions in health research. *Int. J. Med. Inform.* 137:104097. doi: 10.1016/j.ijmedinf.2020.104097

Orchard, L. J., and Fullwood, C. (2010). Current perspectives on personality and internet use. *Soc. Sci. Comput. Rev.* 28, 155–169. doi: 10.1177/0894439309335115

Quan-Haase, A., and Young, A. L. (2010). Uses and gratifications of social media: a comparison of Facebook and instant messaging. *Bull. Sci. Technol. Soc.* 30, 350–361. doi: 10.1177/0270467610380009

Ross, C., Orr, E. S., Sisic, M., Arseneault, J. M., Simmering, M. G., and Orr, R. R. (2009). Personality and motivations associated with Facebook use. *Comput. Human Behav.* 25, 578–586. doi: 10.1016/j.chb.2008.12.024

Rubin, A. M. (1994). "Media uses and effects: a uses-and-gratifications perspective," in *LEA's Communication Series. Media Effects: Advances in Theory and Research*, eds J. Bryant and D. Zillmann (New Jersey: Lawrence Erlbaum Associates), 417–436.

Ruggiero, T. E. (2000). Uses and gratifications theory in the 21st century. *Mass Commun. Soc.* 3, 3–37. doi: 10.1207/S15327825MCS0301_02

Scissors, L., Burke, M., and Wengrovitz, S. (2016). "What's in a Like? Attitudes and behaviors around receiving Likes on Facebook," in *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work and Social Computing - CSCW '16* (New York; NY). doi: 10.1145/2818048.2820066

Vagias, W. M. (2006). *Likert-type scale response anchors*. Clemson International Institute for Tourism and Research Development, Department of Parks, Recreation and Tourism Management, Clemson University.

Whiting, A., and Williams, D. (2013). Why people use social media: a uses and gratifications approach. *Qual. Mark. Res.* 16, 362–369. doi: 10.1108/QMR-06-2013-0041

Yao, R., and Liang, L. (2010). Analysis of the application of simplified NEO-FFI to undergraduates. *Chin. J. Clin. Psychol.* 18, 457–459.

Zywica, J., and Danowski, J. (2008). The faces of Facebookers: investigating social enhancement and social compensation hypotheses; predicting Facebook^{TM} and offline popularity from sociability and self-esteem, and mapping the meanings of popularity with semantic networks. *J. Comput. Mediat. Commun.* 14, 1–34. doi: 10.1111/j.1083-6101.2008.01429.x

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Li and Wang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

APPENDIX

TABLE A1 | A user profile of Liking expressions on WeChat (n = 168) with descriptive statistics of frequency estimates, likelihood (1 = rare, 2 = unlikely, 3 = possible, 4 = likely, 5 = almost certain) and agreement (1 = agree and 5 = disagree) ratings on a 5-point scale.

	Minimum	Maximum	Mean	SD
1. How long in hours do you use WeChat each day?	1	5	2.2	1.2
2. How many WeChat friends do you have?	30	3,000	385.0	322.5
3. How many *like* expressions do you send to other's posts (e.g., text, video, image, article) every 3 days?	1	8	3.3	1.6
4.1. Out of every 10 posts you read on WeChat, how many of them do you agree with?	2.0	9.0	5.7	1.3
4.2. Out of every 10 posts you read on WeChat, how many of them do you think are meaningful?	1.0	9.0	4.4	1.7
4.3. Out of every 10 posts you read on WeChat, how many of them do you trust?	1.0	9.0	6.2	1.6
4.4. Out of every 10 posts you read on WeChat, how many of them do you feel passionate about?	1.0	8.0	4.5	1.6
5.1. How likely are you to click the *like* button to endorse a post you somewhat like?	1.0	5.0	2.9	0.9
5.2. How likely are you to click the like button to endorse a post you like?	1.0	5.0	3.3	0.9
5.3. How likely are you to click the like button to endorse a post you strongly like?	1.0	5.0	4.0	1.0
6.1. How likely are you to click the like button to endorse a post you somewhat dislike?	1.0	5.0	2.1	0.9
6.2. How likely are you to click the like button to endorse a post you dislike?	1.0	5.0	2.3	1.1
6.3. How likely are you to click the like button to endorse a post you strongly dislike?	1.0	5.0	2.8	1.3
7.1. When you agree with a post, how likely are you to click the *like* button?	1.0	5.0	4.1	0.8
7.2. When you agree with a post, how likely are you to make a comment?	1.0	5.0	3.3	1.0
7.3. When you agree with a post, how likely are you to retweet the post to other WeChat friends?	1.0	5.0	2.7	1.1
7.4. When you agree with a post, how likely are you to just keep it to yourself?	1.0	5.0	3.2	1.0
8.1. When you disagree with a post, how likely are you to click the *like* button?	1.0	4.0	1.7	0.8
8.2. When you disagree with a post, how likely are you to make a comment?	1.0	4.0	2.2	1.1
8.3. When you disagree with a post, how likely are you to retweet the post to other WeChat friends?	1.0	4.0	1.8	0.9
8.4. When you disagree with a post, how likely are you to just keep it to yourself?	1.0	5.0	3.8	1.1
8.5. When you disagree with a post, how likely are you to block the author of the post?	1.0	5.0	2.5	1.2
9. Do you think WeChat should add a *Dislike* button?	1	5	2.3	1.1

frontiers in Psychology

ORIGINAL RESEARCH
published: 27 July 2021
doi: 10.3389/fpsyg.2021.664849

Using E-GARCH to Analyze the Impact of Investor Sentiment on Stock Returns Near Stock Market Crashes

Sze Ting Chen[1]* *and Kai Yin Allison Haga*[2]

[1] China-ASEAN International College, Dhurakij Pundit University, Bangkok, Thailand, [2] International Master Program in Asia-Pacific Affairs, College of Social Sciences, National Sun Yat-sen University, Kaohsiung, Taiwan

Purpose: Investor sentiment, the willingness of market participants to invest, is a difficult concept to measure. Exploring the relationship between investor sentiment and stock returns can reveal how investor sentiment affects the operation of the stock market. Such an understanding can assist market participants in making more rational investment decisions based on market laws. Such an understanding can also assist regulators in their roles of supervision and policy making.

Methodology: Although the E-GARCH model has the advantage of considering volatility clustering, it has not previously been used to investigate the impact of investor sentiment changes on the Shanghai Composite Index's market return. This research therefore applies the E-GARCH approach to data from 2015 to 2018, to explore the influence of investor sentiment on the return rate of the Shanghai Composite Index.

Main Findings: There are three main findings. First, when the investor sentiment is increased by the same amount, the rate of return before a stock market crash will have a smaller increase than the rate of change after the crash, which is a new finding. Second, the rate of return on stocks is susceptible to emotional sentiment, rather than simply depending on stock price. Third, the tendency of retail investors to follow the crowd is less in periods of pessimism than it is in periods of optimism, which, in turn, can push up stock yields.

Application: Based on these research results, this article can provide insights to understand how investors' subjective judgments on future earnings affect their investment behavior and how great the impact is on the market. At the same time, it can help investors make more rational investment decisions based on an understanding of market laws, and help regulators with guidance for their supervision and policy making.

Originality/Value: This paper contributes to the theory of the investor sentiment index, improving the index construction method by adding two sentiment proxy indicators: investor activity ACT and stock market leverage level. After constructing the sentiment index and comparing it with the stock market index (Shanghai Composite Index), the fit is found to be improved.

Keywords: investor sentiment index, E-GARCH, small-cap stock effect, yield, financial assets

OPEN ACCESS

Edited by:
Anna Esposito, University of Campania 'Luigi Vanvitelli, Italy

Reviewed by:
Sobia Naseem, Shijiazhuang Tiedao University, China
Mobeen Rehman, Shaheed Zulfiqar Ali Bhutto Institute of Science and Technology, Pakistan

***Correspondence:**
Sze Ting Chen
sze-ting.che@dpu.ac.th

Specialty section:
This article was submitted to Personality and Social Psychology, a section of the journal Frontiers in Psychology

Received: *06 February 2021*
Accepted: *26 May 2021*
Published: *27 July 2021*

Citation:
Chen ST and Haga KYA (2021) Using E-GARCH to Analyze the Impact of Investor Sentiment on Stock Returns Near Stock Market Crashes. Front. Psychol. 12:664849. doi: 10.3389/fpsyg.2021.664849

INTRODUCTION

In July 2020, the "China Securities Journal" pointed out that China's A-share market "is necessary and conditional to enter a healthy bull market," which would be necessary for promoting economic recovery, attracting foreign investment and competing with other countries. Beijing wants to replicate the V-shaped rebound seen in the US stock market, in the hopes that the resulting wealth effect and emotional boost will be effective in solving many domestic economic challenges during the COVID-19 crisis. The current bull market in China is driven by liquidity/policies, with capital flowing from bonds and bank deposits into stocks. Without strong corporate earnings support, market sentiment would tend to be fragile. Moreover, several factors are also potential risks for severe market volatility: disappointing economic data in the second half of the year, corporate earnings in the second quarter and guidance for the second half of the year, geopolitical risks and potential sanctions brought about by Sino-US trade war negotiations and/or political turmoil in Hong Kong, worsening floods in China, and a second wave virus outbreaks (Feng and Liu, 2020). Unlike the 2015 A-share bull market (driven by leverage and margin financing), local investors have thus-far continued to store funds in the market, given that China's domestic bond yields and deposit rates have fallen.

The coronavirus epidemic impacts the real economy. Gong et al. (2020) noted the poor 2020 quarterly reports of some listed companies. Because of the impact of the epidemic on the A-share market, the short-term market has fluctuated sharply before stabilizing. How long this effect will last depends on the evolution of epidemic. From a fundamental point of view, despite the severity of the epidemic, its impact on the real economy is limited, with many stocks having "panic" declines that may reverse in time. Regarding the current valuation level of the A-share market, Chen et al. (2019) argued, even prior to the outbreak of the epidemic, that the current valuation of most individual stocks had reached a historic bottom. The current investor structure in the A-share market is very unbalanced, however, with retail investors accounting for over 90% and institutions accounting for under 10%. Whereas, institutional investment centers on the future growth expectations of listed companies, retail investment tends to be more speculative and influenced by emotion. This large amount of retail investment may amplify swings in the A-share market, with investors first being overly optimistic based hearsay of good performance, and then being panicked by market drops, with a reluctance to reinvest until there is a recovery.

China A shares peaked in the bull market of early 2016 and have yet to regain those heights (Hu and Wang, 2018; Yang, 2018). The sharp rise and fall in 2015 was caused by the Central Bank of China lowering the reserve requirement ratio and the interest rate in April 2014, for the first time in over 2 years. With interest rates dropping, the A-share market rose sharply. Measuring from June 2014, the market would double in less than a year. Such intense speculation created, in turn, fears for the A-share market's outlook. The A-share market began to plummet in mid-June 2015, with multiple crash days for the stock market. The government intervened when the Shanghai Composite Index fell to 4,000 points, but this did not produce a recovery. The market was still far from its bottom, and it was not uncommon to see 1,000 shares fall to their government-imposed daily limit. By 2018, A-shares showed a downward trend throughout the year, owing to factors such as an economic downturn, market volatility, and continued geopolitical uncertainty (e.g., the Brexit process, Sino-US trade disputes and other factors). The Shanghai Composite Index fell 24.59% and the Shenzhen Component Index fell 34.42%, both nearing records with the second largest drops in their histories. In early 2019, the downward trend reversed, but the recovery had stalled by May.

The Chinese capital market is 40 years old (beginning in 1981), but remains immature when compared with the capital markets of developed countries. Whereas, most participants in developed markets are institutional investors (Clark and Monk, 2017), about 70% of Chinese market participants are individual investors, with only 30% being institutional investors. Irrational speculative trading behavior is also more widespread in China's A-share market, with a tendency for investors to chase short-term ups and downs. Also, the information available to the public is very limited, which makes the A-share market more susceptible to rumors. All these factors, as well as others, can easily cause serious deviations in the perception of the market. Therefore, it is particularly important to study the relationship between the stock market and investor sentiment. Past literature has shown that investor sentiment affects stock prices (Renault, 2017; Qadan and Nama, 2018). In subsequent studies, researchers also proposed that investor sentiment affects the value of corporate bonds and options (Bethke et al., 2017; Seok et al., 2019). This research further improves the method of constructing the investor sentiment index by adding two sentiment proxy indicators (investor activity ACT and stock market leverage level) to look at how the investor sentiment fluctuation index is related to stock yield before and after a stock market crash.

Traditional financial theories view investors as generally rational, and consider that rational arbitrageurs can counteract the influence of irrational investors on prices (Daniel and Titman, 1999; Kozak et al., 2018). The core theories of traditional finance include the modern asset portfolio theory of Markowitz (1952), the popular MM theorem proposed by Franco and Merton (1958), the capital asset pricing theory (CAPM) proposed by Sharpe (1964), the efficient market hypothesis proposed by Fama (1970), the option pricing theory proposed by Black and Scholes (1973), and the arbitrage pricing theory (APT) proposed by Ross (2013), etc. Since 1950, traditional finance theories gradually developed into a discipline with a strict logic and unified analytical framework. Since the 1980s, however, financial markets exhibited many anomalies that could not be explained by existing financial theories, such as a momentum effect, an under-reaction anomaly, a small-cap stock over-reaction anomaly, and closed fund discounts behavior.

As a result, behavioral finance theory began to replace traditional finance theories, as it better explains these financial anomalies. Behavioral finance theory employs cognitive psychology and treats investors as being rationally limited in terms of cognition, emotion, and attitude (Day, 2016;

Mushinada, 2020). Investors therefore respond wrongly to market noise and cause systematic cognitive biases. Behavioral finance theory considers how these irrational noise traders affect the operation of efficient markets (De Long et al., 1990; Cuong et al., 2019). Investor sentiment (De Long et al., 1990; Kumari and Mahakud, 2016) is one of the irrational factors that causes noise trading. Studying the relationship between investor sentiment and market returns will therefore contribute to behavioral finance theory. This study can also help investors make more rational investment decisions based on a better understanding of market rules, while also helping regulators in guiding their supervision and policy making.

Research Gap

In recent decades, an increasing number of scholars have studied the impact of investor sentiment on financial asset prices (Zheng, 1999; Kelly and Ahmad, 2018), as well as on investor sentiment and stock prices (Baker and Wurgler, 2006; Smales, 2017). But most of these research works have examined developed countries. Is the situation in the emerging economies consistent with the results for developed countries? Moreover, few scholars have studied China's Shanghai Stock Exchange Index (SHCOMP) before and after stock market crashes. In addition, the SHCOMP is more representative than China's CCTV Watch Index, because the CCTV Watch Index is calculated through a survey of only 60 securities companies and consulting companies in Mainland China. With the continuous development of the capital market, the influence of these 60 institutions in the market will continue to weaken, so it is doubtful whether they can be accurately representative of overall market sentiment. Moreover, In the past, the literature has only talked about the relationship between investor sentiment and stock prices, but it rarely mentioned about whether the relationship between the two before and after the stock market crash was different from usual.

Research Results

This paper not only improves the index construction method by adding two sentiment proxy indicators: investor activity ACT and stock market leverage level but also find that the coefficient before the stock market crash was smaller than that after the stock market crash, in comparing an increase in investor sentiment vs. an equally-sized decrease, the sentiment increase had less effect on the market rate of return before a stock market crash than the decrease had after a crash.

LITERATURE REVIEW AND HYPOTHESES

Research Background

Neoclassical economic theory assumes that people are essentially rational, that information is openly disclosed, and that the market is effectively operated. In other words, people can use available information to rationally invest and consume, and the securities market can respond quickly and accurately to new information. Also, securities prices fully reflect all market information, with market competition making security prices shift from to a new balance in response to independent and random new information.

In the real world, however, the hypothesis of neoclassical economics theory is incomplete. Securities prices cannot fully reflect all the information in the market. New market information may not immediately impact the securities market, nor may the investors' response to that information be rational. Neoclassical economics ignores important aspects of human behavior and psychology. For example, psychological changes affect human decision-making, including investment decisions and consumption behavior. The earliest research on investor sentiment was based on human psychological factors (Watson, 1912). After that, behavioral finance became a field of study that evolved into modern investor sentiment theory. The prospect theory of Kahneman and Tversky (1979) focuses on investor behavior from a psychological perspective, assuming that retail investors are more susceptible to emotional factors or cognitive biases, which lead to unreasonable trading behaviors.

There is no standardized definition of investor sentiment. According to Baker and Wurgler (2006), investor sentiment is defined as "a belief about future cash flows and investment risks that is not justified by the facts at hand." Similarly, Stein (1996) also defined sentiment as a systematic deviation of investors' expectations of the future. Investor sentiment seeks to measure market participants' willingness to invest or expectations for their investments. Although traders experience the undeniable influence of investor sentiment, measuring it presents the challenge of attempting to quantify a feeling (i.e., a sentiment). Moreover, different individual investors may not have the same sentiments. Despite the measurement challenges, it is well-recognized that investor sentiment is a necessary component of stock market behavior, with real explanatory power. In economic activities, sentiment is an uncertain factor, which affects investors' subjective judgments on future returns. This in turn affects its investment behavior and thus has a great impact on the market.

Investor Sentiment and Financial Assets

Early research claimed that noise traders were not important to the formation of financial asset prices (Fama, 1965; Stephen, 1976; Khrennikova, 2016). Talwar et al. (2021) looked at the decision-making process of investors from the perspective of psychology and pointed out that individual investors could easily be affected by emotional factors or cognitive biases, leading to unreasonable trading behaviors. Baker and Wurgler (2006) also proved that investor sentiment could affect stock returns. Behavioral finance theory sees two forces as determining financial asset prices: noise traders and rational arbitrageurs. De Long et al. (1990), Gruber (2011) noted that traders chasing noise may be overactive or inactive, so their behavior could induce stock prices to deviate from their fundamentals. Irrational investors can over-react good news, but they can also ignore under-reported bad news (Brown and Cliff, 2004; Hirota and Sunder, 2016). Shleifer and Vishny (1997) proposed the theory of arbitrage limitation, which argues that the levels of irrational investment behavior and market noise trading are too high to be fully offset by arbitrageurs. As a result, rational arbitrageurs may reduce their positions or withdraw from the market, thereby further weakening the stabilizing effect of arbitrage and causing asset prices to deviate even further from their intrinsic value. Hence,

from the perspective of behavioral finance, investor behavior can be explained by psychological perspectives and bullish/bearish sentiment indexes (Ryu et al., 2017).

Investor sentiment refers to the investment intentions and expectations of investors. Chau et al. (2016) identify such sentiments as the cause behind excessive fluctuations in risk tolerance that are too pessimistic or too optimistic in relation to asset forecasts. Whether pessimistic or optimistic, sentiment can cause an asset's price to deviate from its intrinsic value. The effect is hard to quantify, however, and questions (such as how high this sentiment is or how rapidly it is currently changing) are complicated by the wide differences in interpretations between investors, due to factors ranging from investment style to wealth status. De Long et al. (1990) and Han et al. (2016) proposed a noise trading model, due to the difficulty of predicting the mood of noise traders. They viewed noise trader sentiment as being always present within in the financial market, posing a systemic risk factor that could begin to noticeably affect asset pricing at any time.

Other researchers have tried to understand investor sentiment by combining the ideas of psychology and finance. Ye et al. (2020) noted that sentiment generates when investors rely on their own imperfect understanding of the market and their own cognitive psychology. Brown and Cliff (2004) and Haritha and Uchil (2019) consider investor sentiment as representing the investors' personal expectations for the market's rate of return. For example, investors with high optimism have higher expectations of yield than pessimistic investors have. Baker and Wurgler (2006) and Piccoli and Chaudhury (2018) have linked investor sentiment with a psychological speculative tendency. Rising investor sentiment raises interest in speculative trading, regardless of the differing arbitrage capabilities of the stocks in question. Such speculation therefore has a cross-sectional effect upon the entire market.

To summarize, we agree that modern investors have access to so much market information media that it does influence whether they form bullish or bearish expectations of the stock market's trend in the future. This current research on investor sentiment adopts quantitative methods, using several proxy indicators to approximate investor sentiment. A review of the literature reveals that these indicators fall into two categories: direct indicators of investor sentiment which are also known as an intuitive sentiment indicators, and indirect indicators which are also called objective sentiment indicators.

Investor Sentiment Indicator: Investors Intelligence Index, Bullish Sentiment Index, TURN

Six indicators are primarily used by scholars to measure the intuitive sentiment of investors, particularly in China. (1) The Investor Intelligence Index, or the II Index, has been updated weekly since 1964 and is compiled by Chartcraft Company. Lee et al. (2002) and Jitmaneeroj (2017) have used this index to study investor sentiment. (2) The Friendship Index is compiled by HARDADY. The statistical samples of the index are all institutional investors, including funds, insurance, private equities, and mainstream media. (3) The Haodan Index has been published since 1997 and is the first investor sentiment survey index published in mainland China. (4) The CCTV Watch Index has both daily and weekly frequency data. (5) The Bullish Sentiment Index (BSI) was designed by Cui (2013). (6) The Sina Long Short Index (SinaISI) is compiled by Sina Finance.

Seven proxy indicators of investor sentiment are primarily used by scholars. (1) The closed-end fund discount rate (CEFD) is the most common proxy variable to be used by Western scholars. Neal and Wheatley (1998) regressed the yields of large-cap and small-cap stocks on the New York Stock Exchange (NYSE) from 1933 to 1998 with the discount rate of closed funds over the same period. They found that the discount of closed-end fund could predict the differences between the yields of different types of stocks. When Brown and Cliff (2005) regressed the return rates of large-cap stocks and small-cap stocks over shorter periods from 6 months to 3 years, however, they found that the discount rate of closed funds could not significantly predict the return rate of stocks. (2) The turnover rate (TURN) refers to the frequency of stocks changing hands, as found in the market transaction data. Whereas, trading volume is an absolute quantitative indicator, the turnover rate is a relative quantitative indicator (Lee et al., 2009; Cohen et al., 2016). Baker and Stein (2004) used this indicator to research investor sentiment. They found that when the turnover rate (volume) in the market increases, it is often because the market is more dominated by irrational investors, causing a short-term overvaluation of the market, which indicates that the future rate of return may be low. (3) The Advance-Decline Line Index (ADL) is defined as the ratio of rising stocks (ADV) to declining stocks (DEC) per day, as found in the market performance index. Brown and Cliff (2005) have used this indicator. (4) The financing interest rate belongs to the market leverage index. Financing is the transaction of borrowing funds to buy securities. (5) The securities lending interest rate, which also belong to the market leverage index, is the transaction behavior of borrowing and selling securities. Scholars have used a combination of the financing rate indicator and the securities rate indicator to research investor sentiment (Brown and Cliff, 2004; Chen et al., 2016). (6) The number of new investors is an indicator of investor activity, and represents, to some extent, the views of potential investors who have not yet entered the market (over-the-counter investors). (7) The Investor Sentiment Index is a synthetic index constructed by Zhang et al. (2018). It combines the market turnover rate and closed-end fund discount rate as an objective sentiment indicator.

Nevertheless, when Kim et al. (2014) conducted an empirical study on the Sentiment Index compiled by Investors Intelligence Communications, they concluded that investor sentiment was useless, with no significant impact on market yields in either the short or long term. So more complex models were needed. Brown and Cliff (2005) and Zhang et al. (2018) established a VAR model for investor sentiment and market yield. They found that investor sentiment does have some predictive power for the long-term stock returns, although they agreed with Sot & Statman that this was not the case for short term yields. Baker and Wurgler (2006) and Bekiros et al. (2016) further found that investor sentiment on a cross-section can predict a variety of stock returns. They pointed out that when investor sentiment

was higher (lower), the yields of small-cap stocks, performance-loss stocks, and high volatility stocks in the market were lower (higher), while large-cap stocks, performance-profitable stocks, and low-volatility stocks enjoyed a higher (lower) the rate of return. Engle (2001) studied the dynamic impact of institutional investors on stock returns by constructing a Markov-Switching-GARCH model, and found that an increase in the positions of institutional investors could change the volatile structure of market returns in the short term. In other words, the contribution of institutional investors to market stability was verified.

Investor Sentiment and Stock Yield

Studies have shown that emotions affect stock returns, with investor optimism and pessimism leading to stock price fluctuations (Baker and Wurgler, 2006; Li et al., 2020). Stambaugh et al. (2012) and Renault (2017) found that, in the months following a high-mood period, the long-term anomaly strategy was more profitable than following a low-mood period. Sun et al. (2016) found that, during a recession, investor sentiment had little ability to predict stock returns. Investor sentiment can also change the perception of investors, affect their investment decisions, and lead to changes in stock prices (Stambaugh et al., 2012; Sun et al., 2016).

Cheng and Liu (2005) chose a preference index that represents investor sentiment and then used a VAR model to analyze investor sentiment and stock market returns. The results show that the medium-term investor sentiment index has more influence on the volatility of stock market returns than short-term investor sentiment has. The short-term sentiment index is much more impacted by market fluctuations in the return rate.

An alternative to the VAR model used by Cheng and others is the T-GARCH model used by Fang (2010) and An et al. (2018). The results show that when the market performs poorly, the higher (in a negative sense) is the sentiment index of retail investors, large investors, and institutional investors, and the lower is the daily yield of the Shanghai Stock Exchange. When the market performs well, the better is the investor sentiment and the higher is the daily yield of the Shanghai Composite Index. The T-GARCH model shows that investor sentiment has a good positive predictive ability for stock returns.

Chi et al. (2012) used the Extended Kalman Filter Method to establish their investor sentiment index. They found that investor sentiment has significantly more impact on small-cap stocks than on large-cap stocks. As for the impact of yield on investor sentiment, the yield of large-cap stocks will have a greater impact. Their other finding is that the fluctuations in investor sentiment are more predictive of market returns than the sentiment index itself.

To summarize, after investors receive market information, they will form a bullish or bearish expectation for the market trend. During periods of rising sentiment, investors' expectations for the market outlook will become more optimistic, stimulating high speculative trading and increasing market volume. This will, to a certain extent, promote the overall rise of the whole stock market index. During periods of downward periods of investor sentiment, investors' are more pessimistic about the market outlook, so demand for speculative transactions is weak, which can cause the stock index to drop further. Based on this, we propose our first research hypothesis:

H_1: The more changed is the investor sentiment, the higher the stock market yield

The Leverage Effect commonly discussed in the field of finance produces an asymmetry between positive and negative shocks in the general financial time series. When a company's debt remains unchanged, a lower stock price for the company corresponds to lower shareholder equity. Therefore, the greater the company debt/shareholders' equity, the greater the company's residual risk arising from uncertainty about future cash flow. This effect from a negative shock is greater than for a positive shock. Accordingly, we propose our second research hypothesis:

H_2: The negative impact of investor sentiment on the market yield is greater than the positive impact.

In addition, when studying the correlation between the investor sentiment index and stock returns, some issues have to be further considered. The overall situation of the stock market is always changing, and the influencing factors upon stock return are very complicated. In fact, no small set of factors is able to fully explain the market behavior. In the face of such a volatile market, is the impact of investor sentiment index on stock yields constant? Considering the radical surge and collapse of China's A-shares in 2015, within the sample period studied in this paper, we can also refine and analyze the influence of investor sentiment changes on stock returns before and after that crash. We thus propose our third research hypothesis:

H_3: When increasing investor sentiment is of the same magnitude, the range of increase of the market yield before the stock market crash is smaller than that after the stock market crash.

METHODOLOGY

Sample and Data Collection

The sample data used in this article come from the official website of China Securities Depository and Clearing Corporation Limited (CSDC) and the Wind Financial Terminal. The time span is from January 1, 2015 to December 31, 2018, using all weekly data, for a total of 1,040 sets of data. This time period is selected because we consider that on January 1, 2015, the broad market was continuing a long bear market trend since 2009, where various indexes had downward trends. It was not until April 2015, when the central bank initiated its first RRR cut in over 2 years, that the broad market index began to stabilize. Two more rate cuts occurred in June and November of that year. Then in 2016, the market experienced a sharp spike and falling slide. Finally, by 2018, the market remained relatively stable with no bull or bear trends across the year, although the market had experienced continuous fusing at the beginning of the year. Therefore, by looking at the 4-year market as a whole, the market had experienced a relatively complete bull-bear cycle, which improves the reliability of the analysis and conclusions drawn in this article.

The Selection Process of Emotional Agent Indicators

First, we built an investor sentiment index. Constructing an effective investor sentiment index requires selecting a true proxy index for investor sentiment. Following the standard practice in the literature, we first screened the primary indicators. This screening takes into consideration the possible lead-lag effect of proxy indicators. Baker and Wurgler (2006) indicate that this lead-lag effect of proxy variables of investor sentiment means that these variables can reflect investor sentiment in different periods. We analyze the correlation between the market index and the above-mentioned preliminarily screened sentiment proxy variables and their respective first-order lag variables. Then, we take the variable with the highest correlation coefficient as the corresponding final emotion proxy variable.

To carry out a follow-up empirical study of investor sentiment and stock market returns, this study uses the investor sentiment synthetic index SENTIMENT (Baker and Wurgler, 2006) as the independent variable. The explained variable is the return rate of the stock market. The raw data of the market is selected from the Shanghai Stock Exchange Composite Index (SHCOMP) and the current yield of the Shanghai Stock Exchange (Fang, 2010). This article selects only objective sentiment indicators to represent investors' sentiment in the market, and does not use intuitive indicators.

Intuitive sentiment indicators are generally obtained from specific questions in interviews or questionnaires; the various indexes are then constructed to measure investor sentiment based on the statistical results of investor answers. Although intuitive sentiment indicators can directly measure the feelings and expectations of investors, the validity of such data is still questioned by many industry experts and scholars. There are two main reasons. First, investor ideas may not necessarily be implemented in practice. When investors are interviewed, the views they may express about the future are subject to change. The noise in the market will affect investor decisions, and they may not end up investing in accordance with the opinions expressed in their interviews. Second, it is doubtful whether the sample of an interview survey is representative. For example, the CCTV Watch Market Index is calculated through a survey of 60 securities companies and consulting companies in Mainland China. However, with the continuous development of the capital market, the influence of these 60 institutions in the market will continue to lessen and make them unrepresentative of market sentiment. Based on the above two reasons, we believe that what investors ultimately "do" reflects market sentiment better than what they "think" they will do. In fact, various objective transaction data in the financial market contain the indicators of the sentiment of investors. Hence, this article only uses these objective sentiment indicators.

The construction of our comprehensive indicator of investor sentiment index (SENTIMENT) is as follows. The primary sentiment indicators we have used are market performance indicators (ARMS), market trading indicators (turnover rate, TURN), market activity indicators (investor weekly activity, ACT), and market leverage indicators (financing balance/free float market value ratio of the financing subject, LEVERAGE). The above-mentioned preliminary selected emotional proxy variables and their respective first-order lag variables and market index, were then analyzed for individual correlations. As Baker and Wurgler (2006) noted, these various proxy variables of investor sentiment may each have a different lead-lag effect; we therefore took the above-mentioned preliminary selected emotional proxy variables and their respective first-order lag variables and the market index, and then analyzed them individually for correlations. The variable with the highest correlation coefficient was then taken as the corresponding final emotion proxy variable. Based on these selections, we then use principal component analysis to construct the investor sentiment index of this article.

There are many objective sentiment indicators that characterize investor sentiment. We refer to Fang (2010) to divide the indicators that reflect objective emotions into 4 sub-categories: Market performance indicators (ARMS), Market transaction indicators (turnover rate, TURN), Market Activity Indexes (Investor Weekly Activity, ACT), and Market leverage indicators (the ratio of financing balance to the free market value of the financing subject, LEVERAGE).

Market Performance Indicators: ARMS

In addition to paying attention to the rise and fall (range) of the index after the close of each day, most of the stock investors will take note of how many stocks rose and how many fell each day, in order to perceive the strength and weakness of the stock market performance. One of the most common indicators is the vacancy index (amount of change ratio, ADL), which measures the number of stocks rising and falling each week, with the following ratio formula:

$$ADL_t = \frac{ADV_t}{DEC_t} \tag{1}$$

This ratio measures the number of stocks whose closing prices had risen (ADV_t) or fallen (DEC_t) in week t. When the market sentiment is strong, investor enthusiasm is relatively high and the stock market often shows a general upswing. This upswing means that the ADL ratio will be large. When the market sentiment is down, it will suppress investor enthusiasm to a certain extent. A growing number of declining stocks will lower the ADL ratio. However, the simple measure of the number of rising and falling stocks does not take into account the trading volume of those corresponding stocks (VOLUME). To this end, we make a slight modification to the expression of the Advance-Decline Line (ADL) Index, choosing ARMS as the index to measure market performance, and defining it as:

$$ARMS_t = \frac{ADV_t/VOLUME_ADV_t}{DEC_t/VOLUME_DEC_t} \tag{2}$$

Here, $VOLUME_ADV_t$ represents the transaction amount of rising stocks and $VOLUME_DEC_t$ represents the transaction amount of falling stock prices, in week t.

Market Transaction Indicators: Turnover Rate (TURN)

Turnover rate refers to the frequency with which stocks are bought and sold in the market during a unit of time, so it reflects the liquidity of stocks. When investor sentiment is high, transactional demand leads, driving a race toward stocks that appear likely to generate a quick profit. Therefore, the more active the stock is, the higher the turnover rate should be. In contrast, when investors are depressed, their trading reduces greatly. Their trading behavior will also tend to be conservative, and market transactions will be less, resulting in a low turnover rate. Therefore, the turnover rate index of the stock market can be used to reflect the strength of trading demand and can also be used as an indicator for measuring investor sentiment. We defines the turnover rate as:

$$TURN_t = \frac{\text{Number of traded (Shanghai A Shares}_t + \text{Shenzhen A Shares}_t)}{\text{Free floating market value of (Shanghai A Shares}_t + \text{Shenzhen A Shares}_t)} \quad (3)$$

Market Activity Index: Investor Weekly Activity (ACT)

Regarding the selection of indicators for investor activity, Zhang and Yang (2009) selected the number of newly opened accounts in A-shares. This article uses the Investor Weekly Activity (ACT) indicator. There are two main reasons for this decision. First, measuring A-shares is difficult to changes made by the China Securities Regulatory Commission (CSRC) during our study's test period. The CSRC revised its "Securities Account Business Guide," in April 2015, to allow investors to open A-share accounts at different 20 different brokerage firms; it then revised the policy again the following year, limiting then number to 3. With our sample data covering this time period, these policy changes will result in a distortion of the new investor account opening data. The second reason for not using the A-shares indicator is the problem of unavailable data. At present, the CRSC provides no direct announcement of the investor's weekly activity index. The data that the CSRC does publish (with weekly frequency) are: the number of investors at the end of the period, the number of position holders at the end of the period, and the number of investors who participated in the transaction during the period. Among these three provided data, the number of investors at the end of the period is similar to the number of investors holding positions at the end of the period. We choose to focus on the simple number of investors, because short positions are also a manifestation of attitude. Thus, we define the Investor Weekly Activity (ACT) based on the existing published data:

$$ACT_t = \frac{\text{Number of investors participating in transactions during Period}_t}{\text{Number of investors at the end of Period}_t}. \quad (4)$$

Market Leverage Indicator: Financing Balance/Free Market Value of Financing Subject (Leverage)

In more-developed securities markets, the volume of margin trading and securities lending accounts for a large proportion of the total market transactions. These volumes are thus reflective of how leveraged the market is. Before 2010, China's A-share market only allowed unilateral long positioning. Since 2010, A-shares both long and short positions can be held parallel. A-share margin trading has undergone four large expansions. There are now 948 underlying securities, accounting for about 30% of the total number of A-shares in Shanghai and Shenzhen. However, the current securities lending business is still limited to selected securities firms. Since these firms only have funds in localities and a small amount of stocks, there remain fundamental problems and the growth of this business is slow. Over the same period, only 31 stocks out of 948 double financing target/subjects were securities financing subjects. Among the nearly 940 billion yuan of double financing, securities lending was only 3.5 billion yuan (<0.5%) With this number being essentially negligible, we only consider the impact of the financing balance and define the stock market leverage level (LEVERAGE) as:

$$\text{LEVERAGE}_t = \frac{\text{Financing balance}_t}{\text{Free floating market value of financing subject}_t}.$$

Analytical Procedures

In order to analyze and predict volatile and asymmetric yield fluctuations, this article uses GARCH (1,1) (Bollerslev, 1986) and E-GARCH (Nelson, 1991). These are the two most widely used asymmetric univariate models of conditional volatility for investigating the impact of investor sentiment changes on stock yield. The asymmetric effects on conditional volatility of positive and negative shocks of equal magnitude can be captured in different ways by the exponential GARCH or EGARCH (McAleer and Hafner, 2014). Of GARCH, EGARCH, PGARCH & TGARCH, each model has a divergent purpose with normal error distribution techniques to measure the volatility of investor sentiment, specifically, by using GARCH, ERARCH. Moreover, one of the main advantages of E-GARCH is the model's logarithm of volatility (Mohsin et al., 2019; Salamat et al., 2020). Therefore, during the estimation period, no parameter limitation is required. When estimating the simplest univariate GARCH (1,1), the estimating program usually limits alpha and beta to greater than zero. This is ideal, but in the EGARCH model, such restrictions are not required.

The change of the investor sentiment index is taken as the explanatory variable in the empirical analysis, and the yield sequence of the Shanghai Composite Index is taken

as the explanatory variable. Before modeling the time series and regression analysis, it is generally necessary to perform a stationarity test to avoid a spurious regression. So we conducted a stationarity test to ensure that the time series of the variables in the model are stable. The financial time series data do not have a stable mean, but do tend to be relatively stable in stages that are followed by drastic fluctuations. Simple linear models therefore also need to be tested by ARCH. The ARCH model is a statistical model for time series data. The model describes the variance of the current error term or innovation term as a function of the actual size of the error term in the previous time period.

After the test, we found that the residual sequence does not meet the independence requirement of OLS, so the value estimated by OLS is not unbiased. Therefore, this paper uses GARCH family of models to correct the residuals. The GARCH model requires that all of the coefficients be positive, and that the degree of positive and negative shocks must be consistent. After careful consideration, we selected the E-GARCH model. This model was proposed by Daniel (1991) to solve the problem of asymmetric impacts of positive and negative shocks on market returns in the field of financial analysis. Since it also solves the strict non-negative constraint influence of the GARCH model on parameters, E-GARCH is more suitable for our study. The specific results are shown in **Table 1**.

The "$***$" means significance at the 1% level. It is clear that the market index has a significant relationship with all eight of the above variables, with each correlation coefficient being far above 0. In comparing each variable with its lagged version, two of the four benefit from lagged analysis. The four sentiment indicators selected for this article are therefore LagLEVERAGE, ACT, LagTURN and ARMS. We next use principal component analysis to weigh the four variables and construct the investor sentiment index (SENTIMENT). We implement this process through SAS programming, to obtain the results shown in the following **Table 2**:

According to the results in **Table 3**, we can write the factor expressions of the first and second principal components (1),

$$\begin{aligned} Prin_1 &= 0.532LagLEVERAGE + 0.582ACT + 0.571LagTURN \\ &\quad +0.228ARMS \end{aligned} \tag{5}$$

$$\begin{aligned} Prin_2 &= -0.716LagLEVERAGE - 0.075ACT - 0.146LagTURN \\ &\quad +0.970ARMS \end{aligned} \tag{6}$$

Then, according to the weights of the cumulative contributions of the first two principal components, we weighted the corresponding coefficients in the expressions of $Prin_1$ and $Prin_2$ and obtained the weighted average as the coefficient of the final emotional composite index to arrive at the final sentiment comprehensive index (SENTIMENT). The specific operation rules are as follows.

If we take the first T principal components, the expression of the principal component is $Prin_i = \sum \lambda_{ki} X_i$, $(k \leq \mathrm{T})$. It is also known that the variance contribution degree of the k_{th} principal component is σ_k^2. So the expression of the final emotional synthesis index is:

$$\begin{aligned} \mathrm{Prin} &= \frac{\sum_{k=1}^{T} \lambda_{k1}\sigma_k^2}{\sum_{i=1}^{T} \sigma_i^2} X_1 + \frac{\sum_{k=1}^{T} \lambda_{k2}\sigma_k^2}{\sum_{i=1}^{T} \sigma_i^2} X_2 + \ldots \\ &\quad + \frac{\sum_{k=1}^{T} \lambda_{kn}\sigma_k^2}{\sum_{i=1}^{T} \sigma_i^2} X_n \end{aligned} \tag{7}$$

Referring to (2), we finally get the expression of the investor sentiment index SENTIMENT used in this article:

$$\begin{aligned} SENTIMENT_t &= 0.333LagLEVERAGE_{t-1} + 0.42ACT_t \\ &\quad +0.39LagTURN_{t-1} + 0.436ARMS_t. \end{aligned} \tag{8}$$

Descriptive Statistics

We select the value of investor sentiment innovation (DSENTIMENT) as the explanatory variable of the empirical analysis model, which is define over a period t as:

$$\mathrm{DSENTIMENT_t} = \mathrm{SENTIMENT_t} - \mathrm{SENTIMENT_{t-1}} \tag{9}$$

If $DSENTIMENT_t > 0$, then this means that the newly generated investor sentiment over period t is positive (i.e., new optimism). Similarly a $DSENTIMENT_t < 0$ means that the newly generated investor sentiment is negative (i.e., new pessimism).

The raw data of the explanatory variable is based on the Shanghai Composite Index (SHCOMP), with the current return rate of the Shanghai Composite Index (RetSHCOMP) as the explanatory variable. Also, we use the logarithmic rate of return to express the current rate of return. So the market rate of return in period t is expressed as:

$$\mathrm{RetSHCOMP}_t = ln\mathrm{SHCOMP}_t - ln\mathrm{SHCOMP}_{t-1} \tag{10}$$

The results of descriptive statistical analysis of the market return data are summarized in **Table 4**.

Stationarity Test

In this section, the time series of the explained variables is $\{RetSHCOMP_t\}$. The time series of the explanatory variables is $\{DSENTIMENT_t\}$. Therefore, we first perform unit root tests on these two sequences. This article uses the ADF inspection method. In the ADF test, the null hypothesis and alternative hypothesis are respectively defined as:

$$H_0 : \rho - 1 = 0 \; H_a : \; \rho - 1 < \; 0$$

Here, the null hypothesis H_0 means that the original time series is not stationary. Then in performing the tests of the time series, {Z}, we conduct the following regression:

$$\begin{aligned} Z_t - Z_{t-1} &= (\rho - 1) Z_{t-1} + \beta_0 + \beta_1 t + \beta_2 \Delta Z_{t-1} \\ &\quad +\beta_3 \Delta Z_{t-2} + \nu_t \end{aligned} \tag{11}$$

We implement the ADF stationarity test through SAS. The results of the test for the time series of the explanatory variable

TABLE 1 | The correlation coefficients of the market index SHCOMP and eight primary sentiment indicators.

LEVERAGE	ACT	TURN	ARMS	LagLEVERAGE	LagACT	LagTURN	LagARMS
0.771	0.899	0.811	0.302	0.775	0.891	0.815	0.301
***	***	***	***	***	***	***	***

Source: This research collated.

TABLE 2 | Eigenvalues of the correlation matrix and its cumulative contribution to variance.

	Eigenvalues	Difference	Proportion	Accumulation
1	2.7537	1.8452	0.6885	0.6884
2	0.9086	0.6261	0.2271	0.9156
3	0.2826		0.0706	0.9862

Source: This research collated.

TABLE 3 | Factor loading of the first three principal components.

	Prin1	Prin2	Prin3
LagLEVERAGE	0.533	−0.176	0.826
ACT	0.582	−0.075	−0.332
LagTURN	0.571	−0.146	−0.454
ARMS	0.238	0.970	0.056

Source: This research collated.

TABLE 4 | Basic descriptive statistics of return rate data.

	N	Mean	Standard deviation	Sum	Minimum	Maximum
RetSHCOMP	205	0.0015	0.0349	0.3097	−0.1429	0.0909

TABLE 5 | Unit Root Test Results of RetSHCOMP.

Categories	Lag	Rho	Pr < Rho	Tau	Pr < Tau	*F*	Pr > *F*
Zero Mean	0	−175.04	0.0001	−12.36	<0.0001		
	1	−162.41	0.0001	−8.98	<0.0001		
	2	−160.30	0.0001	−7.51	<0.0001		
Single Mean	0	−175.39	0.0001	−12.35	<0.0001	76.25	0.0001
	1	−163.14	0.0001	−8.98	<0.0001	40.28	0.0001
	2	−161.75	0.0001	−7.53	<0.0001	28.25	0.0001
Trend	0	−175.40	0.0001	−12.32	<0.0001	75.88	0.0001
	1	−163.15	0.0001	−8.95	<0.0001	40.08	0.0001
	2	−161.78	0.0001	−7.50	<0.0001	28.13	0.0001

$RetSHCOMP_t$ and of the explanatory variable $DSENTIMENT_t$ are shown in **Tables 5**, **6**.

Tables 5, **6** show that the time series for the two variables under study in the first question reject the null hypothesis of "H_0: the original series has a unit root (that is, the series is not stationary)." The time series $\{RetSHCOMP_t\}$ and $\{DSENTIMENT_t\}$ are therefore stationary series.

E-GARCH Model Testing

After confirming that these two time series are both stationary, the next step is to establish a linear regression model to examine the impact of investor sentiment changes $\{DSENTIMENT_t\}$ upon the current return rate of the market $\{RetSHCOMP_t\}$. To do so, we first perform the Ljung-Box-Q test on the return sequence $\{RetSHCOMP_t\}$. The test results are shown in **Table 7**.

We can see that none of the Q statistic values can reject the null hypothesis of a white noise sequence. The market return sequence $\{RetSHCOMP_t\}$ is therefore approximately a white noise sequence with no linear autocorrelation. Therefore, we first construct a simple OLS linear model [hereafter referred to as "model (7)"]:

$$\mathrm{RetSHCOMP} = \mathrm{c} + \beta_1 \mathrm{DSENTIMENT_t} + \varepsilon_\mathrm{t} \tag{12}$$

The OLS linear regression model requires that the residual sequence $\{\varepsilon_t\}$ be independent and uniformly distributed. Unfortunately, Financial Time Series usually do not have a stable mean. While most time series are relatively stable in stages, they are often accompanied by severe fluctuations. Because there is an ARCH effect in the stock returns of every bank, the OLS with HAC estimation is not completely correct (Newey and West, 1987). We therefore need to run an ARCH-effect test (Lagranger Multiplier test) on model (7). The judgment index used is the LM statistic. The LM test runs the following regression:

$$\varepsilon_t^2 = \beta_0 + \beta_1 \varepsilon_{t-1}^2 + \beta_2 \varepsilon_{t-2}^2 + \ldots + \beta_m \varepsilon_2^{t-m} + e_t \tag{13}$$

The original hypothesis is H_0, which means that there is no ARCH effect in the residual sequence $\{\ \varepsilon_t\ \}$ up to the order m (that is, $\beta_1 = \beta_2\ = \ldots = \beta_m = 0$) and the alternative hypothesis is H_a(that not all of the β are equal to 0). The LM test results of model (7) are shown in **Table 8** below.

It can be seen that the p-values corresponding to LM statistics up to a 12th order lag are all <1%, which strongly rejects the null hypothesis H_0. This means that the residual sequence $\{\varepsilon_t\}$ does have a high-order ARCH effect. Therefore, it does not meet the independence requirement of the OLS linear regression equation for the residual sequence. We also know that, under the ARCH effect, the β value estimated by OLS linear regression is biased and also ineffective. In order to obtain an unbiased estimate of β, this paper selects the GARCH (1,1) family model to correct the residuals and improve the model (7), in agreement with the procedure of Mohsin et al. (2020) which considers the GARCH (1,1) model to be the best fit model for obtaining volatility.

Assuming that the time series $\{\varepsilon_t\}$ follows the standard GARCH(m,s) model for the time series $\{\varepsilon_t\}$, we obtain the following expression:

$$\begin{cases} \text{Mean equation}: \varepsilon_t = \sqrt{h_t} v_t, \{v_t\}\ \textit{Zero mean white noise sequence} \\ \text{Variance equation}: h_t = a_0 + \sum_{i=1}^{m} a_i \varepsilon_{t-1}^2 + \sum_{j=1}^{s} \beta_j h_{t-j} \end{cases} \tag{14}$$

With the strict constraints of model (9):

$$s.t. \begin{cases} a_0 > 0, a_i \geq 0, \beta_j \geq 0 \\ \sum_{i=1}^{\max(m,s)} a_i + \beta_j < 1 \end{cases} \tag{15}$$

Our research finds that the main constraint of the GARCH model is its higher requirements for the coefficients in the model. Another pair of problems is that the response must be positive and the negative financial shocks must be symmetrical. In fact, in a typical financial time series, negative shocks tend to be stronger than positive shocks. This asymmetry is called the leverage effect in finance. Regarding the issue of the impact of asymmetric shocks, scholars have appropriately extended the traditional standard GARCH model. The more popular extended GARCH models include the T-GARCH model and the E-GARCH model. The E-GARCH model was proposed by Daniel (1996). It not only solves the problem of the asymmetry of positive and negative shocks on market returns, but also solves the constraints that the parameters of the GARCH model be non-negative. Furthermore, The EGARCH model is popular, among other reasons, because it can capture both asymmetry, namely the different effects on conditional volatility of positive and negative effects of equal magnitude, and leverage, which is the negative correlation between returns shocks and subsequent shocks to volatility (McAleer and Hafner, 2014, p. 96.) Mohsin et al. (2020) considered how EGARCH (1,1) assesses the impact of leverage (negative or positive shock) on the unpredictability of bank stock returns. In addition, EGARCH (1,1) uses a logarithmic model of conditional variance, because logarithmic values can be positive or negative, thus avoiding the GARCH model's restriction against non-negative coefficients. We therefore use E-GARCH (1,1). The EGARCH model expression of the current return rate of the market (RetSHCOMP) and investor sentiment changes (DSENTIMENT) are written as:

$$\begin{cases} \text{Mean equation}: RetSHCOMP_t = c + \beta_1 DSENTIMENT_t + \varepsilon_t \\ \text{where } \varepsilon_t = \sqrt{h_t} v_t, \{v_t\}\ \ \textit{Zero mean white noise sequence} \\ \text{Variance equation}: \ln(h_t) = a_0 + a_1 \left| \frac{\varepsilon_{t-1}}{\sqrt{h_{t-1}}} \right| + \phi_{\ 1} \ln\left(h_{t-1}\right) + \theta_{\ 1} \frac{\varepsilon_{t-1}}{\sqrt{h_{t-1}}} \end{cases} \tag{16}$$

Here, if the asymmetric effect coefficient θ_1 is significantly below 0, it indicates that the impact of the shock is a leverage effect. In addition, the impact of the shock is asymmetric if $\theta_1 \neq 0$. Also, there are no non-negative constraints on other parameters in the model.

TABLE 6 | Unit Root Test Results of DSENTIMENT.

Categories	Lag	Rho	Pr < Rho	Tau	Pr < Tau	*F*	Pr > *F*
Zero Mean	0	−199.33	0.0001	−14.09	<0.0001		
	1	−157.72	0.0001	−8.86	<0.0001		
	2	−164.97	0.0001	−7.60	<0.0001		
Single Mean	0	−199.33	0.0001	−14.05	<0.0001	98.71	0.0001
	1	−157.73	0.0001	−8.84	<0.0001	39.10	0.0001
	2	−164.98	0.0001	−7.58	<0.0001	28.75	0.0001
Trend	0	−199.62	0.0001	−14.03	<0.0001	98.39	0.0001
	1	−158.64	0.0001	−8.85	<0.0001	39.14	0.0001
	2	−166.38	0.0001	−7.58	<0.0001	28.76	0.0001

TABLE 7 | {$RetSHCOMP_t$} serial white noise autocorrelation test results.

Lag	*Q* statistics	Degree of freedom	Pr > *Q*	Autocorrelation coefficient					
6	5.20	6	0.518	0.140	0.052	0.013	0.026	0.031	−0.027
12	13.67	12	0.323	−0.001	0.004	0.096	−0.065	−0.149	0.057
18	15.07	18	0.657	−0.050	−0.035	0.000	0.048	0.016	0.006
24	23.55	24	0.488	0.081	0.029	0.168	−0.004	−0.004	−0.034

TABLE 8 | Model (7) ARCH test of residuals.

Order	*Q*	Pr > *Q*	LM	Pr > LM	Order	*Q*	Pr > *Q*	LM	Pr > LM
1	8.58	0.0034	8.52	0.0035	7	103.22	<0.0001	60.75	<0.0001
2	42.37	<0.0001	36.58	<0.0001	8	115.30	<0.0001	60.74	<0.0001
3	48.92	<0.0001	37.16	<0.0001	9	126.20	<0.0001	62.90	<0.0001
4	69.28	<0.0001	41.49	<0.0001	10	130.49	<0.0001	66.38	<0.0001
5	69.27	<0.0001	46.02	<0.0001	11	133.37	<0.0001	66.55	<0.0001
6	95.21	<0.0001	56.22	<0.0001	12	134.22	<0.0001	66.93	<0.0001

TABLE 9 | Parameter estimation results of model (11) (1).

Variable	Estimate	Standard Error	*t* value	Approximate Pr > |t|
C	0.0026	0.0018	1.44	0.1492
Q	0.0252	0.0063	4.07	<0.0001
a_0	−0.1231	0.0791	−1.56	0.1109
a_1	0.1418	0.0754	1.88	0.0601
ϕ_1	0.9824	0.0112	88.32	<0.0001
θ_1	−1.1901	0.7014	−1.70	0.0897

TABLE 10 | Parameter estimation results of model (11) (2).

SSE	Log likelihood	AIC	Total R party	Normality test	Pr > Chi-square
0.2360	441.4553	−870.9104	0.0436	1.8229	0.4020

RESULTS

After choosing the E-GARCH (1,1) model as the modification scheme of model (7), and the final regression results areas **Tables 9**, **10**:

The above results produce three observations. First, the EGARCH model did not reject the null hypothesis that the residual sequence is white noise, which supports that the model estimation is well-realized. Second, in the mean value equation, the β coefficient is significantly above 0 throughout the sample

time, which shows that there is a positive relationship between the market return (RetSHCOMP) and changes in investor sentiment (DSENTIMENT). Therefore, when investor sentiment generates new optimism (pessimism), the market rate of return tends to correspondingly increase (decrease). This proves that our hypothesis H_1 does not reject the null hypothesis, and is supported. Third, it can also be seen from the parameter estimation results of Model 11 that the asymmetric effect coefficient θ_1 is significantly below 0. This shows that, over the entire time period from 2015 to 2018, the impact of negative sentiment on the market return is greater than the impact of positive sentiment, indicating that there is a leverage effect in the market. As in Kasman's et al. (2011) research, EGARCH assessed the impact of leverage (positive or negative shock) on the unpredictability of bank stock returns. This proves that hypothesis H_2 is supported, with the negative impact of investor sentiment on the market return being greater than the positive impact.

Sectional Inspection Results Before and After the Stock Market Crash

The input data from 2015 contain a sharp spike and crash in China's A-share market. In order to analyze the data in better detail, we divide the sample data (from 2015 to 2018) into two time periods: before and after the stock market crash on June 12, 2015. Then, we build an E-GARCH (1,1) model of the market return sequence $\{RetSHCOMP_t\}$ on the investor sentiment change sequence $\{DSENTIMENT_t\}$ based on the (1,1) model for the above two time periods. The final parameter estimation results from the model are shown in **Table 11**.

From the results shown in **Table 11**, the β coefficient is significantly >0, both before and after the stock market crash. This shows that the market return rate had a positive correlation with the changes in investor sentiment. When there is a positive (negative) investor sentiment with new interest, the market return rate tends to increase (decrease). This result is the same as the estimated result of the whole stage. Also, we found that the β coefficient before the stock market crash was significantly smaller than that after the crash. This means that, when investor sentiment is increased (reduced) by the same magnitude, the rate of increase (decrease) in the market rate of return before the stock market crash is smaller than the rate of change in the market rate of return after the stock crash. This proves that hypothesis H_3 is supported.

CONCLUSION AND DISCUSSION

In China's stock market, fluctuations in the investor sentiment index have a profound impact on stock market yields. To a certain extent, this means that the effectiveness of the Chinese securities market is insufficient, with many other factors affecting the operation of the market. The reasons for this may include the imperfect system of China's securities market (Xie, 2016; Zhanga and Yaob, 2016), the need for stronger supervision of regulatory authorities, and the professionalism level of the investment participants. China's capital market has developed rapidly in recent years, expanding its market value, transaction volume, and company listings. If compared with the securities markets of developed countries in the West, however, the problems and deficiencies in China's securities markets are clear. Be it from the perspective of laws and regulations, of market systems, or of market efficiency, it is precisely because of these shortcomings that investor sentiment has such a significant impact on yields in China's securities market.

Studying various factors that affect stock returns, predicting and measuring returns of various types of financial products, and exploring the relationship between risk and return – all these have become important topics in finance. Concerning modern investment portfolio theory, classic financial theories lack explanatory power for many real-world stock price fluctuations. The academic community should strengthen its analysis of the many practical problems in the financial market, from the perspective of behavioral finance theory in emerging economies, as well as strengthening the completeness of capital asset theory and systematic development (Cassetti et al., 2020). This would help promote the development of the entire financial discipline. Focusing on the impact of investor sentiment on the yield of assets can guide improvements to the market and to the investment philosophy of investors. Policymakers, investors, and researchers are most interested in understanding stock returns and market performance, and how it interacts with return predictors. The Chinese government wants more foreign capital flow and wants to understand its impact on the market (Howes et al., 2017). As an emerging economy with unstable policies, with a fast changing environment, and with markets that have not yet been fully developed, there is a wide range of explanatory variables that may affect Chinese stock returns.

In addition to combining the investor sentiment index research of many prior scholars, this article further constructs a new investor sentiment index that adds two sentiment proxy indicators: investor activity ACT and stock market leverage level LEVERAGE. The final constructed sentiment index is found to produce a better fit than the broader market index (Shanghai Composite Index). We established the above-mentioned E-GARCH(1,1) model on the return sequence and on investor sentiment index changes to conduct an empirical analysis of this issue. Our study found that investor sentiment, whether optimistic or pessimistic, has a significant impact on stock market returns, indicating that investor sentiment fluctuation is one of the factors affecting stock market price trends.

Contributions

This study offers several contributions to financial researchers. Although previous researchers have explored the relationships between investor sentiment and its various predictors on the basis of a cross-sectional design (Zhang et al., 2017; Ren et al., 2018), this study investigates the two emotional proxy indicators: investor activity (ACT) and stock market leverage level (LEVERAGE). Unlike previous studies that measured investor sentiment at a specific point by the movement of capital between markets (Ding et al., 2017; Lan et al., 2020), this study considers China's securities market and finds that the investor sentiment rate has a more marked impact. To some

TABLE 11 | Parameter estimation results of the model (11) before and after the stock market crash.

	Before the stock market crash			After the stock market crash		
Variable	**Estimate**	***t*-value**	**Approximate Pr > |t|**	**Estimate**	***t*-value**	**Approximate Pr > |t|**
C	0.0032	1.33	0.1845	−0.0023	−0.63	0.5273
□	0.0191	2.57	0.0102	0.0478	3.61	0.0003
a_0	−0.2362	−1.29	0.1973	−1.7703	−6.1	<0.0001
a_1	0.0042	0.04	0.9693	0.4221	1.4	0.1619
ϕ_1	0.9671	40.37	<0.0001	0.7415	16.76	<0.0001
θ_1	−1.4972	−1.81	0.0783	−1.0444	−1.05	0.2932

extent, this means that the effectiveness of China's securities market is insufficient, and there are still many other factors that affect the operation of the market. In view of the fact that the influence of investor sentiment in China's securities market is widespread, the steady and effective development of China's securities market requires securities market regulators to improve laws and supervision (Abbasi and Riaz, 2016). Also, it is important to educate investors on investment concepts. Furthermore, the investor sentiment index constructed in this article provides a risk control perspective.

This research also finds that, although the traditional capital asset pricing model has made a great contribution to market pricing, it is still imperfect, with many of its assumptions being questioned by various empirical studies. First of all, in the field of asset pricing, we believe that, to expand finance theory, behavioral finance and traditional finance need to continue to merge, and the asset pricing framework needs to adapt to the different characteristics of different capital markets. As for China's securities market, continuous improvement and maturity are inevitable, but the speed of the process is vitally related to governmental reform, introducing institutional laws, and improving the education of investors.

Managerial Implications

This study has implications to the academic community and subsequent researchers. First, studying various factors that affect stock returns, predicting and measuring returns of various types of financial products, and exploring the specific relationship between risk and return are important topics of research. Starting from modern portfolio theory, various models have been proposed to study the risks and returns of financial assets (Guironnet et al., 2016; Li et al., 2018). However, with the development of classic financial theories and the maturity of the capital market, we find that the research results of classic financial theories lack explanatory power for many actual conditions of stock price fluctuations in the market. To bridge the gap between theory and practice, this article recommends that, in addition to the development of classic theories, the academic community should strengthen the analysis of the many practical problems in modeling financial markets. This analysis can consider the perspective of emerging behavioral finance theories, and can promote the integration and systems of capital asset theory development. In any market, as long as the investors are people, psychology and investor sentiment will continue to impact the rate of return.

Second, for market participants, a better understanding of the impact of investor sentimental fluctuations can encourage investors to rationally analyze the market. At present, China's securities market is dominated by private investors without professional, capital and information advantages. Such investors are more likely to blindly speculate on short-term up-and-down trends. Compared with professional investment institutions, their market analysis ability and risk tolerance are low. It is better to encourage investors to make medium- and long-term value investments to reduce irrational behaviors in the investment process (Antony, 2020; Zhang, 2020). Therefore, in the securities market, adding institutional investors would help to stabilize the market structure. Improving the overall investment quality of market traders can reduce the number of noise traders (Peress and Schmidt, 2020) and bring stock prices closer to their actual values.

Third, from the supervising agency perspective, these agencies also use various methods to guide investment participants to establish a good investment philosophy and enhancing investors' emotional awareness and financial knowledge. Whether an institutional investor or an individual investor, all investors should always distinguish between different industries and companies, and between short, medium and long-term investments, so as to treat each differently and accordingly. Investors should focus on the intrinsic value of investment stocks, should be vigilant about risks, and should not one-sidedly pursue high returns while ignoring risks. Individual investors, especially, should treat the market rationally, strive to overcome human weaknesses, and use more comprehensive risk management tools to control risks and their own positions.

Limitation and Future Research Recommendations

This work has a number of limitations which suggest opportunities for future research. First, because the research budget and time were constrained, this study used only four emotional agent indicators to test the research model. Since the Chinese securities market is an immature emerging market, future studies should collect more data to obtain adequate information about what differences exist between the behavioral financial theory of emerging economies and the traditional

financial theory of Western established markets. Second, this article only considers the impact of changes in investor sentiment on stock market returns. In the actual operation of a stock market, changes in stock market returns may in turn affect investor sentiment. This impact should be two-way and should be explored further in the future. Third, this article does not consider the macroeconomic cycle variables to predict and measure the yield of various types of financial products. Nor does it explore the specific relationship between risk and return. Fourth, investor sentiment changes over time according to environmental conditions, making it difficult to determine which specific stocks attract speculators or arbitrageurs. Fifth, this research did not discuss the rationality and ability of stock investors to choose stocks. Since some scholars have proposed that investors have the ability to choose stocks (Li et al., 2016), subsequent researchers still need to further distinguish between smart investors and retail investors.

From the empirical results of this article, we find that foreign capital plays an important role in both the spot market and the futures market in China. Therefore, we recommend that governing authorities should hasten opening up to foreign investment, in order encourage the domestic financial market to align with international standards. This paper can also be used as a reference for the opening up of financial policies in emerging markets. In addition, the empirical results, which have explored the operation methods of foreign investment in periods of high and low sentiment for the futures and spot markets, can be used as a reference for domestic institutional legal entities and general investors. In addition, changes in investor sentiment are a systemic factor affecting stock returns. Investor sentiment thus has a strong ability to predict the volatility of future stock returns. This article also found that investor sentiment in China's Shanghai Stock Exchange (SHCOMP) stock market will significantly affect stock price volatility, with the impact of investor sentiment on volatility being principally through the channel of rewards.

Here are the suggestions for follow-up studies. (1) This study has used the foreign-invested sentiment proxy variables to screen out four indicators. We suggest that technical analysis indicators can be added into these sentiment indicators to see whether it could be more explanatory. (2) Concerning the impact of foreign sentiment indicators on the excess returns and volatility of futures and spot goods, this article only studies the overall market. We suggest that future research efforts should further divide the market into long and short terms, so that the impact of the sentiment indicator on the long market and the short markets can be analyzed separately, while observing the level of impact when the sentiment is extremely high or extremely low. (3) The constructed sentiment indicator can be further used as a predictive model to test whether it has predictive ability for the excess returns in the futures and spot markets.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

REFERENCES

Abbasi, F., and Riaz, K. (2016). CO2 emissions and financial development in an emerging economy: an augmented VAR approach. *Energy Policy* 90, 102–114. doi: 10.1016/j.enpol.2015.12.017

An, N., Wang, B., Pan, P., Guo, K., and Sun, Y. (2018). Study on the influence mechanism of air quality on stock market yield and volatility: empirical test from China based on GARCH model. *Finance Res. Lett.* 26, 119–125. doi: 10.1016/j.frl.2017.12.002

Antony, A. (2020). Behavioral finance and portfolio management: review of theory and literature. *J. Public Aff.* 20:e1996. doi: 10.1002/pa.1996

Baker, M., and Stein, J. (2004). Marker liquidity as a sentiment indicator. *J. Finan. Markets* 7, 271–299. doi: 10.1016/j.finmar.2003.11.005

Baker, M., and Wurgler, J. (2006). Investor sentiment and the cross-secion of stock returns. *J. Finance* 61, 1645–1680. doi: 10.1111/j.1540-6261.2006.00885.x

Bekiros, S., Gupta, R., and Kyei, C. (2016). A non-linear approach for predicting stock returns and volatility with the use of investor sentiment indices. *Appl. Econ.* 48, 2895–2898. doi: 10.1080/00036846.2015.1130793

Bethke, S., Gehde-Trapp, M., and Kempf, A. (2017). Investor sentiment, flight-to-quality, and corporate bond comovement. *J. Banking Finance* 82, 112–132. doi: 10.1016/j.jbankfin.2017.02.007

Black, F., and Scholes, M. (1973). The pricing of options and corporate liabilities. *J. Polit. Econ.* 81, 637–654. doi: 10.1086/260062

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity, *J. Econometr.* 31, 307–327. doi: 10.1016/0304-4076(86)90063-1

Brown, G., and Cliff, M. (2004). Investor sentiment and the near-term stock market. *J. Empir. Finance*11, 1–27. doi: 10.1016/j.jempfin.2002.12.001

Brown, G., and Cliff, M. (2005). Investor sentiment and asset valuation. *J. Bus.* 78, 405–440. doi: 10.1086/427633

Cassetti, V., Powell, K., Barnes, A., and Sanders, T. (2020). A systematic scoping review of asset-based approaches to promote health in communities: development of a framework. *Glob. Health Promot.* 27, 15–23. doi: 10.1177/1757975919848925

Chau, F., Deesomsak, R., and Koutmos, D. (2016). Does investor sentiment really matter? *Int. Rev. Finan. Anal.* 48, 221–232. doi: 10.1016/j.irfa.2016.10.003

Chen, J., Kadapakkam, P. R., and Yang, T. (2016). Short selling, margin trading, and the incorporation of new information into prices. *Int. Rev. Finan. Anal.* 44, 1–17. doi: 10.1016/j.irfa.2016.01.002

Chen, R., Yu, J., Jin, C., and Bao, W. (2019). Internet finance investor sentiment and return comovement. *Pacific-Basin Finan. J.* 56, 151–161. doi: 10.1016/j.pacfin.2019.05.010

Cheng, K, and Liu, R. (2005). Research on the interaction of investor sentiment and stock market. *Shanghai Econ. Res.* 11:86–93. doi: 10.3969/j.issn.1005-1309.2005.11.012.10.3969

Chi, L., Zhang, G., Zhuang, X., and Song, D. (2012). Investor sentiment indicators and stock markets: a study based on the extended Kalman filter method. *J. Manage. Eng.* 126, 132–169. doi: 10.1080/13504851.2011.577003

Clark, G. L., and Monk, A. H. (2017). *Institutional Investors in Global Markets*. Oxford: Oxford University Press. doi: 10.1093/oso/9780198793212.001.0001

Cohen, G., Blake, R. S., and Goodman, D. (2016). Does turnover intention matter? Evaluating the usefulness of turnover intention rate as a predictor of actual turnover rate. *Rev. Pub. Person. Administr.* 36, 240–263. doi: 10.1177/0734371X15581850

Cui, C. (2013). Investor sentiment measurement method. *Management Research*. Available online at: https://wenku.baidu.com/view/50302b0add36a32d737581be (accessed June 14, 2021).

Cuong, P. K., Ngoc, T. T. B., Cong, B. T., and Chau, V. T. Q. (2019). Noise trader risk: evidence from vietnam stock market. *Hue Univ. J. Sci. Econ. Dev*. 128, 5–16. doi: 10.26459/hueuni-jed.v128i5C.5083

Daniel, B. M. (1996). Conditional heteroskedasticity in asset returns: a new approach. *Model. Stock Market Volatil.* 59, 37–64. doi: 10.1016/B978-012598275-7.50004-1

Daniel, K., and Titman, S. (1999). Market efficiency in an irrational world. *Finan. Anal. J.* 55, 28–40. doi: 10.2469/faj.v55.n6.2312

Daniel, N. B. (1991). Conditional Heteroskedasticity in Asset Returns: a newapproach. *Econometrica - The Econometric Soc.* 59, 347–370. Available online at: http://links.jstor.org/sici?sici=0012-9682%28199103%2959%3A2%3C347%3ACHIARA%3E2.0.CO%3B2-

Day, G. (2016). Irrational investors and the corporate inversion puzzle. *SMUL Rev.* 69:453. Available online at: http://digitalrepository.smu.edu (accessed June 14, 2021).

De Long, J. B., Shleifer, A., Summers, L. H., and Waldmann, R. J. (1990). Noise trader risk in financial markets. *J. Polit. Econ.* 98, 703–738. doi: 10.1086/261703

Ding, Z., Liu, Z., Zhang, Y., and Long, R. (2017). The contagion effect of international crude oil price fluctuations on Chinese stock market investor sentiment. *Appl. Energy* 187, 27–36. doi: 10.1016/j.apenergy.2016.11.037

Engle, R. (2001). GARCH 101: the use of ARCH/GARCH models in applied econometrics. *J. Econ. Perspect.* 15, 157–168. doi: 10.1257/jep.15.4.157

Fama, E. F. (1965). The behavior of stock-market prices. *J. Bus.* 38, 34–105. doi: 10.1086/294743

Fama, E. F. (1970). Efficient capital market:a review of theory and empirical work. *J. Finance* 25, 383–417. doi: 10.2307/2325486

Fang, Y. (2010). An Empirical Study on the Impact of Securities Investor Sentiment on Market Returns and Volatility. *Business Times* 004, 94–96.

Feng, Z., and Liu, D. (2020). China and Hong Kong stock market views: Chinese technology stocks – timely technical adjustments. *Market Strategy*.

Franco, M., and Merton, H. M. (1958). The cost of capital, corporation finance and the theory of investment. *Am. Econ. Rev.* 48, 261–297.

Gong, B., Zhang, S., Yuan, L., and Chen, K. Z. (2020). A balance act: minimizing economic loss while controlling novel coronavirus pneumonia. *J. Chin. Govern.* 5, 249–268. doi: 10.1080/23812346.2020.1741940

Gruber, M. J. (2011). Another puzzle: the growth in actively managed mutual funds. *J. Finance* 51, 117–144. doi: 10.1111/j.1540-6261.1996.tb02707.x

Guironnet, A., Attuyer, K., and Halbert, L. (2016). Building cities on financial assets: the financialisation of property markets and its implications for city governments in the Paris city-region. *Urban Stud.* 53, 1442–1464. doi: 10.1177/0042098015576474

Han, B., Tang, Y., and Yang, L. (2016). Public information and uninformed trading: implications for market liquidity and price efficiency. *J. Econ. Theory* 163, 604–643. doi: 10.1016/j.jet.2016.02.012

Haritha, P. H., and Uchil, R. (2019). Impact of investor sentiment on decision-making in Indian stock market: an empirical analysis. *J. Adv. Manage. Res*. 17, 66–83. doi: 10.1108/JAMR-03-2019-0041

Hirota, S., and Sunder, S. (2016). "Price bubbles sans dividend anchors: evidence from laboratory stock markets," in *Behavioral Interactions, Markets, and Economic Dynamics*, eds S. Ikeda, H. Kato, F. Ohtake, and Y. Tsutsui (Tokyo: Springer), 357–395. doi: 10.1007/978-4-431-55501-8_13

Howes, M., Wortley, L., Potts, R., Dedekorkut-Howes, A., Serrao-Neumann, S., Davidson, J., et al. (2017). Environmental sustainability: a case of policy implementation failure? *Sustainability* 9:165. doi: 10.3390/su9020165

Hu, J. P., and Wang, J. (2018). *Chinese capital market: an empirical overview*. NBER Working Paper No. 24346. National Bureau of Economic Research, Cambridge, MA, United States. doi: 10.3386/w24346

Jitmaneeroj, B. (2017). Does investor sentiment affect price-earnings ratios? *Stud. Econ. Finance* 34, 183–193. doi: 10.1108/SEF-09-2015-0229

Kahneman, D., and Tversky, A. (1979). On the interpretation of intuitive probability: a reply to Jonathan Cohen. *Cognition* 7, 409–411. doi: 10.1016/0010-0277(79)90024-6

Kasman, S., Vardar, G., and Tunç, G. (2011). The impact of interest rate and exchange rate volatility on banks' stock returns and volatility: evidence from Turkey. *Econ. Model.* 28, 1328–1334. doi: 10.1016/j.econmod.2011.01.015

Kelly, S., and Ahmad, K. (2018). Estimating the impact of domain-specific news sentiment on financial assets. *Knowledge-Based Syst.* 150, 116–126. doi: 10.1016/j.knosys.2018.03.004

Khrennikova, P. (2016). Application of quantum master equation for long-term prognosis of asset-prices. *Physica A* 450, 253–263. doi: 10.1016/j.physa.2015.12.135

Kim, J. S., Ryu, D., and Seo, S. W. (2014). Investor sentiment and return predictability of disagreement. *J. Bank. Finance* 42, 166–178. doi: 10.1016/j.jbankfin.2014.01.017

Kozak, S., Nagel, S., and Santosh, S. (2018). Interpreting factor models. *J. Finance* 73, 1183–1223. doi: 10.1111/jofi.12612

Kumari, J., and Mahakud, J. (2016). Investor sentiment and stock market volatility: evidence from India. *J. Asia-Pacific Bus.* 17, 173–202. doi: 10.1080/10599231.2016.1166024

Lan, Y., Huang, Y., and Yan, C. (2020). Investor sentiment and stock price: empirical evidence from Chinese SEOs. *Econ. Model.* 94, 703–714. doi: 10.1016/j.econmod.2020.02.012

Lee, T. Y., Tzeng, W. C., Lin, C. H., and Yeh, M. L. (2009). Effects of a preceptorship programme on turnover rate, cost, quality and professional development. *J. Clin. Nurs.* 18, 1217–1225. doi: 10.1111/j.1365-2702.2008.02662.x

Lee, W., Jiang, C., and Indro, D. (2002). Stock market volatility, excess returns, and the role of investor sentiment. *J. Bank. Finance* 26, 2277–2299. doi: 10.1016/S0378-4266(01)00202-3

Li, M., Li, W., Wang, F., Jia, X., and Rui, G. (2020). Applying BERT to analyze investor sentiment in stock market. *Neural Comput. Appl.* 33, 4663–4676. doi: 10.1007/s00521-020-05411-7

Li, Z., Dong, H., Huang, Z., and Failler, P. (2018). Asymmetric effects on risks of Virtual Financial Assets (VFAs) in different regimes: a case of Bitcoin. *Quan. Finan. Econ.* 2, 860–883. doi: 10.3934/QFE.2018.4.860

Li, Z., Wang, F., and Dong, X. (2016). Are all investment decisions to subscribe to new stocks mindless?: investor heterogeneity and behavior in the process of subscribing to new stocks. *China J. Account. Res.* 9, 283–304. doi: 10.1016/j.cjar.2016.09.002

Markowitz, H. (1952). Portfolio selection. *J. Finance* 7, 77–91. doi: 10.1111/j.1540-6261.1952.tb01525.x

McAleer, M., and Hafner, C. M. (2014). A one line derivation of EGARCH. *Econometrics* 2, 92–97. doi: 10.3390/econometrics2020092

Mohsin, M., Naiwen, L., Zia-UR-Rehman, M., Naseem, S., and Baig, S. A. (2020). The volatility of bank stock prices and macroeconomic fundamentals in the Pakistani context: an application of GARCH and EGARCH models. *Oeconomia Copernicana* 11, 609–636. doi: 10.24136/oc.2020.025

Mohsin, M., Naseem, S., Muneer, D., and Salamat, S. (2019). The volatility of exchange rate using GARCH type models with normal distribution: evidence from Pakistan. *Pac. Bus. Rev. Int.* 11, 124–129.

Mushinada, V. N. C. (2020). Are individual investors irrational or adaptive to market dynamics? *J. Behav. Exp. Finance* 25:100243. doi: 10.1016/j.jbef.2019.100243

Neal, R., and Wheatley, S. (1998). Do measures of investor sentiment predict returns? *J. Finan. Anal.* 33, 523–547. doi: 10.2307/2331130

Nelson, B. D. (1991). Conditional heteroskedasticity in asset returns: a new approach, *Econometrica* 59, 347–370. doi: 10.2307/2938260

Newey, W., and West, K. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica* 55, 703–708. doi: 10.2307/1913610

Peress, J., and Schmidt, D. (2020). Glued to the TV: distracted noise traders and stock market liquidity. *J. Finance* 75, 1083–1133. doi: 10.1111/jofi,.12863

Piccoli, P., and Chaudhury, M. (2018). Overreaction to extreme market events and investor sentiment. *J. Appl. Econ. Lett.* 25:52. doi: 10.1080/13504851.2017.1302052

Qadan, M., and Nama, H. (2018). Investor sentiment and the price of oil. *Energy Econ.* 69, 42–58. doi: 10.1016/j.eneco.2017.10.035

Ren, R., Wu, D. D., and Liu, T. (2018). Forecasting stock market movement direction using sentiment analysis and support vector machine. *IEEE Syst. J.* 13, 760–770. doi: 10.1109/JSYST.2018.2794462

Renault, T. (2017). Intraday online investor sentiment and return patterns in the US stock market. *J. Bank. Finance* 84, 25–40. doi: 10.1016/j.jbankfin.2017.07.002

Ross, S. A. (2013). "The arbitrage theory of capital asset pricing," in *Handbook of the Fundamentals of Financial Decision Making: Part I*, eds L. C. MacLean and W. T. Ziemba (Hackensack, NJ: World Scientific Publishing Co. Pte. Ltd), 11–30. doi: 10.1142/9789814417358_0001

Ryu, D., Kim, H., and Yang, H. (2017). Investor sentiment, trading behavior and stock returns. *Appl. Econ. Lett.* 24, 826–830. doi: 10.1080/13504851.2016.1231890

Salamat, S., Lixia, N., Naseem, S., Mohsin, M., Zia-ur-Rehman, M., and Baig, S. A. (2020). Modeling cryptocurrencies volatility using GARCH models: a comparison based on Normal and Student's T-Error distribution. *Entrepreneurship Sustain. Issues* 7, 1580–1596. doi: 10.9770/jesi.2020.7.3(11)

Seok, S. I., Cho, H., and Ryu, D. (2019). Firm-specific investor sentiment and daily stock returns. *North Am. J. Econo. Finance* 50:100857. doi: 10.1016/j.najef.2018.10.005

Sharpe, W. F. (1964). Capital asset prices: a theory of market equilibrium under conditions of risk. *J. Finance* 19, 425–442. doi: 10.1111/j.1540-6261.1964.tb02865.x

Shleifer, A., and Vishny, R. W. (1997). The limits of arbitrage. *J. Finance* 52, 35–55. doi: 10.1111/j.1540-6261.1997.tb03807.x

Smales, L. A. (2017). The importance of fear: investor sentiment and stock market returns. *Appl. Econ.* 49, 3395–3421. doi: 10.1080/00036846.2016.1259754

Stambaugh, R. F., Yu, J., and Yuan, Y. (2012). The short of it: investor sentiment and anomalies. *J. Finan. Econ.* 104, 288–302. doi: 10.1016/j.jfineco.2011.12.001

Stein, J. C. (1996). Rational capital budgeting in an irrational world. *J. Bus.* 69, 429–455. doi: 10.1086/209699

Stephen, A. R. (1976). The arbitrage theory of capital asset pricing. *J. Econ. Theory* 13, 341–360. doi: 10.1016/0022-0531(76)90046-6

Sun, L., Najand, M., and Shen, J. (2016). Stock return predictability and investor sentiment: a high-frequency perspective. *J. Bank. Finance* 73, 147–164. doi: 10.1016/j.jbankfin.2016.09.010

Talwar, M., Talwar, S., Kaur, P., Tripathy, N., and Dhir, A. (2021). Has financial attitude impacted the trading activity of retail investors during the COVID-19 pandemic? *J. Retail. Consum. Serv.* 58:102341. doi: 10.1016/j.jretconser.2020.102341

Watson, G. N. (1912). VII. A theory of asymptotic series. *Philos. Transac. Royal Soc. London A* 211, 279–313. doi: 10.1098/rsta.1912.0007

Xie, Y. (2016). "Research on the feasibility of enterprise IPO registration system in China," in *2016 International Conference on Economics, Social Science, Arts, Education and Management Engineering* (Huhhot: Atlantis Press). doi: 10.2991/essaeme-16.2016.183

Yang, D. (2018, July 16). A-shares are at the bottom, attracting continued inflow of foreign capital. *Shanghai Securities News*. Available online at: http://finance.eastmoney.com/a/201807169069106 16.html (accessed June 14, 2021).

Ye, J., Li, D., and Cao, Y. (2020). Investor irrational selection bias in stock market based on cognitive psychology: evidence from herding behaviour. *Rev. Argentina Clín. Psicol.* 29:90. doi: 10.24205/0327616.2020.13

Zhang, G., Wang, J., Guo, H., and Zhang, X. (2018). "The relationship between investor sentiment and stock market volatility: based on the VAR model," in Wuhan International Conference on e-Business. Association For Information Systems (Wuhan).

Zhang, Q., and Yang, S. (2009). Noise trading, investor sentiment fluctuations and stock returns. *Syst. Eng. Theory Pract.* 29, 40–47. doi: 10.1016/S1874-8651(10)60010-5

Zhang, W. (2020). Application of crowd thought and herd behavior in economic investment. *Revista Argentina Clín Psicol.* 29, 328–333. doi: 10.24205/03276716.2020.44

Zhang, Y., Zhang, Y., Shen, D., and Zhang, W. (2017). Investor sentiment and stock returns: evidence from provincial TV audience rating in China. *Physica A* 466, 288–294. doi: 10.1016/j.physa.2016.09.043

Zhanga, S., and Yaob, Y. (2016). A research on stock price manipulation in China. *Eur. J. Bus. Econ. Account.* 4, 72–81.

Zheng, L. (1999). Is money smart? A study of mutual fund investors' fund selection ability. *J. Finance* 54, 901–933. doi: 10.1111/0022-1082.00131

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Chen and Haga. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Psychology

BRIEF RESEARCH REPORT
published: 02 August 2021
doi: 10.3389/fpsyg.2021.684357

Emotional Gaze: The Effects of Gaze Direction on the Perception of Facial Emotions

Jing Liang[1†], Yu-Qing Zou[2†], Si-Yi Liang[3], Yu-Wei Wu[4] and Wen-Jing Yan[3,5]*

[1] School of Educational Science, Ludong University, Yantai, China, [2] Faculty of Psychology, Southwest University, Chongqing, China, [3] College of Teacher Education, Wenzhou University, Wenzhou, China, [4] Wenzhou Business College, Wenzhou, China, [5] School of Mental Health, Wenzhou Medical University, Wenzhou, China

Previous research has found that when gaze direction matches the underlying behavioral intent communicated by the expression of a specific emotion, it enhances or facilitates the perception of that emotion; this is called the shared signal hypothesis (SSH). Specifically, a direct gaze shares an approach-orientated signal with the emotions of anger and joy, whereas an averted gaze shares an avoidance-orientated signal with fear and sadness. In this research, we attempted to verify the SSH by using different materials on Asian participants. In Experiment 1 we employed photos of models exhibiting direct and averted gazes for rating tasks, in order to study the effects of gaze direction on participants' perception of emotion. In Experiment 2 we utilized smiling faces in a similar investigation. The results show that for neutral and smiling faces, a direct gaze (relative to a gaze of avoidance) increased the likelihood of a subject perceiving a happy mood; a gaze of avoidance increased the likelihood that anger and fear would be perceived. The effect of gaze direction on emotional expression perception was verified, but a "facilitating-impairing" pattern was not. The difference between our work and previous research may be attributable to the materials employed (which were more ecological), as well as the participants, who were from a different culture.

Keywords: gaze, emotion perception, facial expression, culture, shared signal theory

OPEN ACCESS

Edited by:
Anna Esposito, University of Campania 'Luigi Vanvitelli, Italy

Reviewed by:
Anne Böckler, Julius Maximilian University of Würzburg, Germany
Jie Zhang, Jiangnan University, China

***Correspondence:**
Wen-Jing Yan
yanwj@wzu.edu.cn

† These authors have contributed equally to this work

Specialty section:
This article was submitted to Emotion Science, a section of the journal Frontiers in Psychology

Received: *23 March 2021*
Accepted: *15 June 2021*
Published: *02 August 2021*

Citation:
Liang J, Zou Y-Q, Liang S-Y, Wu Y-W and Yan W-J (2021) Emotional Gaze: The Effects of Gaze Direction on the Perception of Facial Emotions. Front. Psychol. 12:684357. doi: 10.3389/fpsyg.2021.684357

INTRODUCTION

Among all nonverbal signals, gaze is one of the most attractive facial features and conveys much information (Admoni and Scassellati, 2017). Perceiving others' eyes serves distinct social and emotional functions. A gaze usually indicates interest (approach-avoidance), since people often look at things they like and avoid things they don't (Shimojo et al., 2003; Bayliss et al., 2007). Moreover, gaze can have a significant influence on emotional expression perception. For example, angry people often stare into the eyes of the person with whom they are trying to quarrel or fight, and timid people who fear others may drop their eyes and look away. Previous research has found that when gaze direction matches the underlying behavioral intent (approach-avoidance) communicated by an emotional expression, the perception of that emotion is enhanced or facilitated (Adams and Kleck, 2003, 2005).

These findings can be explained by the shared signal hypothesis (SSH) (Rigato et al., 2013). The SSH proposes that when different cues convey the same information (such as approaching or avoidance), attributions can be made with increased certainty, facilitating the processing of

either cue. Interference can occur when different cues convey conflicting information, affecting the processing of each. It has been argued that proximity-oriented emotions such as happiness, love, and anger tend to be expressed through direct vision, while avoidance-oriented emotions such as jealousy, sadness, and disgust are more likely to be communicated by avoidance (Argyle and Cook, 1976; Kleinke, 1986; Fehr and Exline, 1987). For example, angry faces are detected faster and perceived with greater intensity when the expresser's gaze is directed toward the observer rather than away, while an averted gaze enhances the perception and detection of fearful faces (Rigato et al., 2013). In addition, the SSH may be valid from infancy. Previous work has shown that both adults and infants have demonstrated the ability to discern approach- and avoidance-oriented emotions, matching them with direct and averted gazes, respectively.

In Adams and Kleck (2003) research, it was observed that angry facial expressions were more quickly decoded (i.e., there was less reaction latency in recognition) when displayed in conjunction with a direct rather than averted gaze; in contrast, fear expressions were more quickly decoded when displayed in conjunction with an averted rather than direct gaze. Joyful expressions were more quickly decoded when displayed with a direct rather than averted gaze, and sadness expressions were more quickly decoded when displayed with an averted rather than direct gaze. In Adams and Kleck (2005) work, the SSH was further supported by three studies. The researchers found that a direct gaze led to more angry and joyful disposition attributions, whereas an averted gaze led to more fearful and sad disposition attributions. Benton (2010) examined reactions to briefly presented direct and averted faces displaying expressions of fear and anger; the results supported the notion of signal congruence as a mechanism through which gaze and viewpoint affect our responses to facial expressions. These are very interesting findings and provide a more profound understanding of the SSH and the relationship between gaze direction and emotional expression.

However, the work of Adams and Kleck (2003) faced certain challenges. Bindemann et al. (2008) conducted six experiments to re-examine their results. The latter study indicated that the perception of emotional expression was impaired in a speed-based classification task when the eyes of the face stimulus were averted. In rating tasks, the results were not incongruent, indicating that the perception of selected expressions enhanced under an averted gaze was stimulus was task-bound. Overall, the findings of Adams and Kleck (2003) seemed not to be robust, and instead were situationally dependent. Bindemann et al. (2008) believed it was unresolved how approach/avoidance theory mapped onto existing psychological data, including in previous studies conducted with fearful faces. There is extensive evidence that the perception of fear is intimately linked with a direct eye gaze, characterized by a wide scleral contrast above the iris that disappears as the eyes are turned sideways.

Importantly, the notion that an angry expression involves staring directly at another is not universal, especially taking cultural differences into consideration. Culture has the power to shape not only one's worldview, but also broad psychological functions such as perception, attention, and cognition (Markus and Kitayama, 1991; Varnum et al., 2010). In many East Asian contexts, direct eye contact can be perceived as threatening and disrespectful, and an averted eye gaze is more appropriate; in many Western contexts the reverse is true. For example, angry averted gazes may be interpreted as an attempt to mitigate the negative emotional arousal of the target and maintain the relationship (Kitayama et al., 2006; Park et al., 2018). Moreover, an averted gaze may show dislike and withdrawal from the intention to attack, and thus is a more suitable way to indicate anger in many situations. From this perspective, facial expressions with an averted gaze are more likely to be perceived as angry in Asian cultures, according to the SSH. Therefore, we are curious about whether such a variety of styles of the SSH for the emotion-gaze relationship also exist for Chinese people, who put a much greater premium on collectivism than individualism (Oyserman et al., 2002).

In addition to reexamining the study by Adams and Kleck (2005), the present research tested the effects of gaze on smiling faces. Smiles are usually considered an approachable facial expression and indication of happiness. However, many psychologists agree that smiles can reflect a vast array of emotions, rather than a simple universal expression of happiness. People may smile when they're frightened, flirtatious, horrified, or mortified (Ambadar et al., 2009). The emotion associated with a smile is easily affected by gaze direction (e.g., an embarrassed smile reveals itself through an averted gaze). As previous findings on neutral faces have shown, smiling faces may also be more likely to be rated higher for sadness and fear, or such a mixture may generate other emotions.

The manipulation of gaze type (i.e., direct vs. averted) can be accomplished by using an image processing tool such as Photoshop. Such operations provide strict control because only the gaze direction is changed. However, a manually crafted gaze is not particularly natural and sometimes can look strange. As seen in Figure 1 in Ganel (2011), the eyes seemed to be devoid of spirit and may have exerted an unwanted influence over the observers. As Bindemann et al. (2008) claimed, the advantage of "real gaze" stimuli is the reflection of natural variations in eye gaze direction, while the disadvantage is that the stimuli may vary slightly in ways other than gaze, even within the same identity.

In this research, we studied how gaze affects people's emotional rating of neutral and smiling faces. In Experiment 1 we asked the models to display both direct and averted gazes and investigated whether previous findings held true in different situations. We further investigated smiles with direct and averted gazes in Experiment 2 to explore how the direction of the gaze changed the perception of a smile.

EXPERIMENT 1

Method

The procedure received ethical approval from the IRB at the university, and written informed consent was obtained from all participants.

To acquire photos showing averted and direct gazes in both smiling and neutral faces, we recruited 42 college students (21

males and 21 females) and took their pictures; these served as stimuli for the experiment. Participants read and signed consent forms and agreed that their photos could be used for scientific research. Two cameras were mounted on tripods at a height of 120 cm and placed in front of each participant. Individuals were asked to sit on a chair, hold a remote control, look at the middle camera positioned straight ahead of them (with their nose pointed toward the middle camera), and press the remote control to take the photo. We asked participants to display both neutral and smiling faces. To ensure that the participants were smiling naturally, we tried to amuse them by providing recordings of laughter or teasing them. During the entire procedure, the faces were recorded in the full-frontal position. All of the photographs were cropped in order to display only the head and neck of each individual. We copied and mirrored right-gaze photos to obtain left-gaze pictures. In our experiment, each direct-gaze photo was presented twice to balance out the design. A model displaying both left- and right-averted gazes could balance the "direction effect" (if there is any) and generate more stable results for each model in the four trials. All were resized to a standard 320 pixels in width, with variable heights to preserve the aspect ratio (see **Figure 1**). In this experiment, we used 60 photos from 20 models randomly.

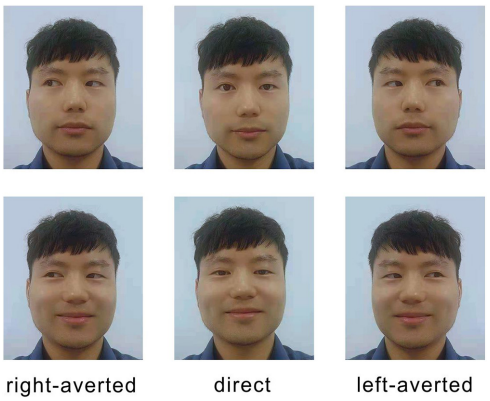


FIGURE 1 | Demo of the materials used in the experiment. The first and second rows indicate the neutral and smiling faces, respectively.

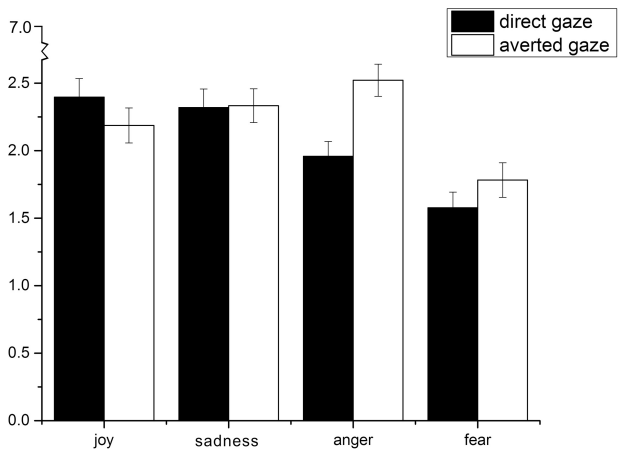


FIGURE 2 | Effects of gaze direction of neutral faces on perceived emotion dispositions. It should be noted that the *Y*-axis indicates a continuous rating scale, which actually ranged from 1 (not at all frequently) to 7 (very frequently) for the task.

Participants and Experiment Apparatus

A power analysis by G*Power[1] 3.1.9.21 indicated $N = 36$ to detect an effect size of 0.25, which is a large effect, with a repeated-measures analysis of variance (ANOVA) and within-factors test (*F*-test), with a number of groups one and number of measurements four, with a probability of $1–\beta = 0.95$, $\alpha = 0.05$. Assuming possibly invalid or missing data, we recruited more participants than were required. A total of 44 undergraduate students (age: 22.71 ± 1.38 years; 27 females) participated in this study as raters of face photos. All signed consent forms. We used computers with 21-inch LCD monitors (resolution at 1024 × 768 pixels) and employed the software package E-prime 2.0 for stimulus presentation and data collection.

Design and Procedure

Each participant was seated at a distance of approximately 70 cm from the presentation screen at separate tables in the lab. The researchers verbally explained the participants' task to make sure they understood how to proceed. All faces showed neutral displays and were presented in random order. Participants were asked to rate each of 80 faces on four emotion scales (i.e., anger, fear, sadness, and joy). There was a 1-min break after 40 trials. The instructions were as follows: "Please rate each of the faces on four emotion scales (anger, fear, sadness, and joy) and enter the associated number in the input box. Ratings are made on a seven-point continuous scale ranging from 1 (not at all frequently) to seven (very frequently)." Each trial began with a fixation cross at the center of the screen that lasted for a duration of 500 ms. This was then replaced by one of the stimulus faces, which remained on the screen until a response was made. Ratings were made on a seven-point continuous scale ranging from 1 (not at all frequently) to seven (very frequently). The next trial began immediately after the preceding response was made.

Results

In order to test the predicted interaction between gaze direction and perceived emotion disposition, a four (anger/joy/fear/sadness emotion dimension) × 2 (direct/averted gaze direction) repeated measures ANOVA was conducted. A main effect for emotion was found, $F(3,129) = 11.565$, $p < 0.001$, $\eta_p^2 = 0.212$, indicating that the emotion rating was different between types, regardless of gaze direction (see **Figure 2**). The post hoc test showed that faces were rated lower for the fear disposition than for other emotion dispositions (fear = 1.68, anger = 2.24, sadness = 2.33, joy = 2.29). A main effect also emerged for gaze, $F(1,43) = 32.428$, $p < 0.001$,

[1]G*Power is a tool used to compute statistical power analyses. G*Power can also be used to compute effect sizes and graphically display the results of power analyses. http://stats.idre.ucla.edu/other/gpower/

η_p^2 = 0.430, such that faces showing an averted gaze were rated higher overall for the likelihood to experience emotion (M = 2.21, SE = 0.091) than were faces showing a direct gaze (M = 2.06, SE = 0.101).

Consistent with our expectations, an emotion-gaze interaction was found, $F(3,129) = 40.315$, $p < 0.001$, $\eta_p^2 = 0.484$, indicating that the direction of the gaze resulted in different effects on perceptions of emotion. A simple effects analysis was conducted to assess these effects for each emotion condition. As predicted, a direct gaze (M = 2.40, SE = 0.92), relative to an averted gaze (M = 2.19, SE = 0.87), increased the perceived likelihood of the stimulus person being perceived as having a joyful disposition, $t(43) = 6.125$, $p < 0.01$, $r = 0.969$. Conversely, an averted gaze (M = 2.51, SE = 0.80), relative to a direct gaze (M = 1.94, SE = 0.72), was more often associated with an angry disposition, $t(43) = -8.953$, $p < 0.001$, $r = 0.851$. Likewise, an averted gaze (M = 1.72, SE = 0.76), relative to a direct gaze (M = 1.51, SE = 0.62), was more often associated with a fearful disposition, $t(43) = -3.707$, $p = 0.001$, $r = 0.901$.

Discussion

This study found that gaze direction systematically influenced the perceived emotion disposition conveyed by a neutral face. In this experiment, a direct gaze was attributed to a more joyful disposition, whereas an averted gaze was attributed to an angrier or more fearful disposition. However, Adams and Kleck found that a direct gaze enhanced the perception of approach (i.e., anger and joy) and improved the mood of approach (i.e., anger and joy), while avoiding eye contact enhanced the perception of avoidance (i.e., fear and sadness). In this experiment, as compared to an averted gaze, a direct gaze increased the subject's perception of a happy mood; an averted gaze increased the perception of anger and fear relative to a direct gaze. The differences in our findings could be due to variations in the materials used in the experiment, as well as cultural incongruities. These will be discussed in the general discussion below.

The next step was to study whether gaze direction would affect the perception of emotion in smiling expressions. Smiling is an approaching signal, but when it is accompanied by an averted gaze, it is more likely to be judged as sadness or fear, as previous findings on neutral faces have found. Alternatively, the sense of anger and fear might be blocked by the smile. In Experiment 2 we changed the intensity of the facial expressions, using a smiling expression to determine whether doing so would affect the role of gaze direction in emotion perception.

EXPERIMENT 2

In the pre-experiment, we randomly presented neutral and smiling faces to participants. However, the results showed that the neutral and smiling faces influenced one another. The smiling faces exhibited strong emotions, which made the neutral faces overwhelmingly more likely to be rated as neutral. Therefore, in Experiment 2 we only presented smiling faces.

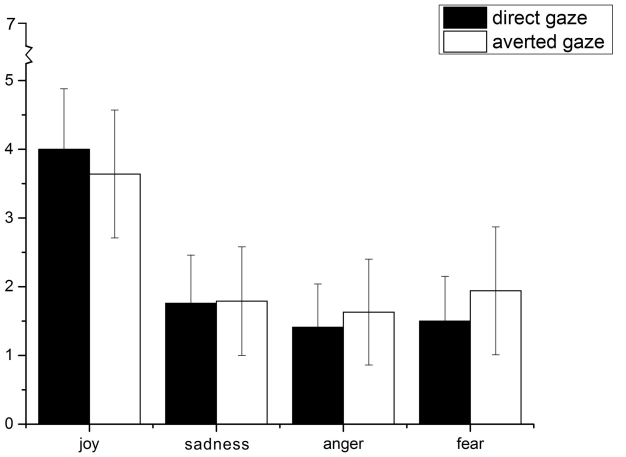


FIGURE 3 | Effects of gaze direction of smiling faces on perceived emotions. It should be noted that the Y-axis indicates a continuous rating scale that ranged from 1 (not at all frequently) to 7 (very frequently) for the task.

Method

A total of 28 female and nine male undergraduate students participated in this study. Individuals were asked to rate the facial displays regarding the emotion expressed, using an emotion profile comprized of a number of scales (i.e., happiness, anger, fear, and sadness) ranging from 1 (not at all frequently) to 7 (very frequently). Except for all faces displaying positive emotions, all other conditions were the same as in Experiment 1.

Results

In order to test the predicted interaction between gaze direction and perceived emotion disposition, a 4 (anger/joy/fear/sadness emotion dimension) × 2 (direct/averted gaze direction) repeated measures ANOVA was computed. A main effect for emotion was found, $F(3,108) = 132.016$, $p < 0.001$, $\eta_p^2 = 0.786$. Inspection of the means indicated that this effect was due to the faces being rated higher with regards to the joyful disposition than the other emotion dispositions (joy = 3.80, sadness = 1.77, anger = 1.52, fear = 1.72). A main effect also emerged for gaze, $F(1,36) = 8.614$, $p < 0.05$, $\eta_p^2 = 0.193$, such that averted gaze faces were rated higher overall with regards to the likelihood of experiencing emotion (M = 2.25, SE = 0.101) than were direct gaze faces (M = 2.19, SE = 0.089) (see **Figure 3**). Consistent with our expectations, these main effects were qualified by an emotion/gaze direction interaction, $F(3,108) = 16.388$, $p < 0.001$, $\eta_p^2 = 0.313$.

Direct t-tests were then conducted to assess the reliability of these effects for each emotion condition. As predicted, the direct gaze (M = 4.00, SE = 0.88), relative to the averted gaze (M = 3.64, SE = 0.93), increased the likelihood of the stimulus person being perceived as having a joyful disposition, $t(36) = 4.26$, $p < 0.001$. Conversely, an averted gaze (M = 1.63, SE = 0.77), relative to a direct gaze (M = 1.41, SE = 0.63), was more often associated with an angry disposition, $t(36) = -3.132$, $p < 0.05$. Likewise,

an averted gaze (M = 1.94, SE = 0.93), relative to a direct gaze (M = 1.50, SE = 0.65), was more often associated with a fearful disposition, $t(36) = -4.588$, $p < 0.001$.

GENERAL DISCUSSION

This study found that gaze direction affected people's perceptions of emotion in neutral expressions, but the pattern identified was different from that uncovered in previous research. Adams and Kleck (2005) found that the direction of the eyes' gaze affected the emotions perceived, depending on the specific type of emotion; a direct gaze could improve the mood of approach (i.e., anger and joy), while a gaze avoiding direct eye contact could enhance the perception of avoidance (i.e., fear and sadness). In Experiment 1 a direct gaze, relative to a gaze of avoidance, increased the likelihood of a subject perceiving a happy mood; a gaze of avoidance increased the likelihood that anger and fear would be perceived. Neutral faces had no emotional tendency, and thus were easily affected by gaze direction. However, when we changed the emotion type in Experiment 2 by using smiling faces, we also found that the perception of emotion (usually positive emotions) could be significantly modulated by gaze direction within the same pattern. Eye gaze could modulate the perception of emotion not only in neutral faces, but also for other types of facial expressions.

However, in our study, the "facilitating-impairing" pattern displayed was different from that found in Adams and Kleck (2005). One reason for this may be that we used different materials and designs. The materials used in Adams and Kleck (2005) is different in several aspects: (1) half the faces were manipulated to display direct gaze and half averted gaze; (2) gaze direction was manipulated using Adobe Photoshop; (3) if an exemplar face was presented with direct gaze in one stimulus set, it was presented with averted gaze in the other set. In our study, the "within-stimulus" approach was chosen to better control for individual differences because we considered that the approach-avoidance feeling is largely affected by the model's appearance. Moreover, in the present study, gaze direction was directly provided by participants. This operation was the same as what was used in Bindemann et al. (2008), but their results showed that real and manually crafted gazes produced similar results. Averting the eyes may also affect other facial muscles and somewhat distort general facial expressions. For example, when participants tried to avert their eyes without adjusting their head orientation, they felt unnatural and not as comfortable as usual. Our pattern was also different from what was found in Bindemann et al. (2008), where an averted gaze impaired overall emotional expression perception. They used both real and manually crafted gazes and obtained similar results. In addition, compared with the design in Adams and Kleck (2005), the participants in the present study were more likely to remember their earlier ratings, consequently increasing the similarity of the ratings for direct and averted gazes. The results show that even though there might have been such an impact, the effect size was still large. This design proves that the effect of gaze direction is very robust from another point of view.

In addition, cultural differences may have been responsible for this incongruity. Researchers have noted that individualism and independent cultural expressions emphasize the direct and explicit communication of emotions (Markus and Kitayama, 1991), while in collectivist cultures such as those of East Asia, the suppression of emotions is more encouraged than is emotional expression (Soto et al., 2011). In fact, in Western cultures, people tend to practice independent self-interpretation; denying oneself emotional expression and experience is equivalent to denying one's true self. Conversely, in Asian countries such as Japan, China, and South Korea, people are more collectivist and interdependent. Controlling and suppressing emotional expressions is considered key to maintaining a harmonious relationship with the group (Markus and Kitayama, 1991); direct expressions of emotion have a negative impact on interpersonal and collective relationships. Anger and an averted gaze seem to be culturally congruent in these countries. This is partly supported by Park et al. (2018), where P1 in the Angry-Averted condition was significantly larger than P1 in the Neutral-Averted condition ($p < 0.005$). The conclusion was that Asian Americans allocated more early attentional processing to averted angry eyes, as compared to neutral and direct angry gazes. The authors explained that averted gazes may reflect a culturally appropriate emotional attunement and could be interpreted as an attempt to mitigate the negative emotional arousal of the target and an attempt to maintain the relationship. Therefore, in Eastern cultures, people tend to avert their eyes to reduce the negative impact of interpersonal communication; thus, averted eyes are more likely to be perceived as angry. From this perspective, the SSH is also supported by our research, though the pattern may be different due to the different "shared signal pattern." In the current study, a direct gaze shared an approach-orientated signal with joy, whereas an averted gaze shared an avoidance-orientated signal with fear and anger.

There are some limitations in the present study. We used only neutral and smile faces as the materials. It is likely that the spontaneity of the smiles could be a potential factor that affect the emotion perception and the elicited smiles in the present study may vary in spontaneity. In addition, in the future studies different facial expression should be further considered. As for the manually crafted gaze, it should be compared with the real directly in the future study and the results could be more persuasive when the differences were explained by cultural differences.

This study used ecological experiment materials to study the effects of eye orientation on the perception of emotion by Chinese subjects. The SSH and effect of gaze direction on emotional expression perception were verified for neural and smiling faces, but the SSH pattern could be affected by various factors. Whether such differences are due to

the materials, procedure, or culture is a topic requiring further investigation.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by IRB in Wenzhou University. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

W-JY and JL conceived and designed the experiments. Y-WW and Y-QZ performed the experiments. Y-QZ, JL, S-YL, and W-JY analyzed the data. JL, Y-QZ, S-YL, and W-JY wrote and revised the manuscript.

FUNDING

This research was supported by the National Natural Science Foundation of China (No. 31900761), Natural Science Foundation of Zhejiang Province (No. LQ16C090002), Shandong Provincial Natural Science Foundation (ZR2018PC034), and A Project of Shandong Province Higher Educational Science and Technology Program (J17RB179).

REFERENCES

Adams, R. B. Jr., and Kleck, R. E. (2003). Perceived gaze direction and the processing of facial displays of emotion. *Psychol. Sci.* 14, 644–647. doi: 10.1046/j.0956-7976.2003.psci_1479.x

Adams, R. B. Jr., and Kleck, R. E. (2005). Effects of direct and averted gaze on the perception of facially communicated emotion. *Emotion* 5, 3–11. doi: 10.1037/1528-3542.5.1.3

Admoni, H., and Scassellati, B. (2017). Social eye gaze in human-robot interaction: a review. *J. Hum. Robot Interact.* 6, 25–63. doi: 10.5898/jhri.6.1.admoni

Ambadar, Z., Cohn, J. F., and Reed, L. I. (2009). All smiles are not created equal: morphology and timing of smiles perceived as amused, polite, and embarrassed/nervous. *J. Nonverbal Behav.* 33, 17–34. doi: 10.1007/s10919-008-0059-5

Argyle, M., and Cook, M. (1976). *Gaze and Mutual Gaze*. Cambridge: Cambridge U Press.

Bayliss, A. P., Frischen, A., Fenske, M. J., and Tipper, S. P. (2007). Affective evaluations of objects are influenced by observed gaze direction and emotional expression. *Cognition* 104, 644–653. doi: 10.1016/j.cognition.2006.07.012

Benton, C. P. (2010). Rapid reactions to direct and averted facial expressions of fear and anger. *Vis. Cogn.* 18, 1298–1319. doi: 10.1080/13506285.2010.481874

Bindemann, M., Burton, A. M., Leuthold, H., and Schweinberger, S. R. (2008). Brain potential correlates of face recognition: geometric distortions and the N250r brain response to stimulus repetitions. *Psychophysiology* 45, 535–544. doi: 10.1111/j.1469-8986.2008.00663.x

Fehr, B. J., and Exline, R. V. (1987). "Social visual interaction: a conceptual and literature review," in *Nonverbal Behavior and Communication*, eds A. W. Siegman and S. Feldstein (Mahwah, NJ: Lawrence Erlbaum Associates, Inc), 225–325.

Ganel, T. (2011). Revisiting the relationship between the processing of gaze direction and the processing of facial expression. *J. Exp. Psychol. Hum. Percept. Perform.* 37, 48–57. doi: 10.1037/a0019962

Kitayama, S., Mesquita, B., and Karasawa, M. (2006). Cultural affordances and emotional experience: socially engaging and disengaging emotions in Japan and the United States. *J. Pers. Soc. Psychol.* 91, 890–903. doi: 10.1037/0022-3514.91.5.890

Kleinke, C. L. (1986). Gaze and eye contact: a research review. *Psychol. Bull.* 100, 78–100. doi: 10.1037/0033-2909.100.1.78

Markus, H. R., and Kitayama, S. (1991). Culture and the self: implications for cognition, emotion, and motivation. *Psychol. Rev.* 98, 224–253. doi: 10.1037/0033-295x.98.2.224

Oyserman, D., Coon, H. M., and Kemmelmeier, M. (2002). Rethinking individualism and collectivism: evaluation of theoretical assumptions and meta-analyses. *Psychol. Bull.* 128, 3–72. doi: 10.1037/0033-2909.128.1.3

Park, G., Lewis, R. S., Wang, Y. C., Cho, H. J., and Goto, S. G. (2018). Are you mad at me? Social anxiety and early visual processing of anger and gaze among Asian American biculturals. *Cult. Brain* 6, 151–170. doi: 10.1007/s40167-018-0067-1

Rigato, S., Menon, E., Farroni, T., and Johnson, M. H. (2013). The shared signal hypothesis: effects of emotion-gaze congruency in infant and adult visual preferences. *Br. J. Dev. Psychol.* 31, 15–29. doi: 10.1111/j.2044-835x.2011.02069.x

Shimojo, S., Simion, C., Shimojo, E., and Scheier, C. (2003). Gaze bias both reflects and influences preference. *Nat. Neurosci.* 6, 1317–1322. doi: 10.1038/nn1150

Soto, J. A., Perez, C. R., Kim, Y.-H., Lee, E. A., and Minnick, M. R. (2011). Is expressive suppression always associated with poorer psychological functioning? A cross-cultural comparison between European Americans and Hong Kong Chinese. *Emotion* 11, 1450–1455. doi: 10.1037/a0023340

Varnum, M. E., Grossmann, I., Kitayama, S., and Nisbett, R. E. (2010). The origin of cultural differences in cognition: the social orientation hypothesis. *Curr. Dir. Psychol. Sci.* 19, 9–13. doi: 10.1177/0963721409359301

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Liang, Zou, Liang, Wu and Yan. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Psychology

BRIEF RESEARCH REPORT
published: 01 September 2021
doi: 10.3389/fpsyg.2021.683259

The Humble Charisma of a White-Dressed Man in a Desert Place: Pope Francis' Communicative Style in the Covid-19 Pandemic

Rosa Scardigno, Concetta Papapicco, Valentina Luccarelli, Altomare Enza Zagaria, Giuseppe Mininni and Francesca D'Errico*

Department of Educational Sciences, Psychology, Communication, University of Bari "A. Moro", Bari, Italy

The context of deep uncertainty, fear, and "social distancing" characterizing the COVID-19 pandemic has led to a need for cultural anchorages and charismatic leaders who may conjointly and effectively support human beings, strengthen their identity, and empower social commitment. In this perspective, the charismatic leadership of Pope Francis, which is widely shared not only within the religious world, may play a crucial role in facing emergency with existential reasons and psychological resources. The general aim of this work is to shed light on the communicative features of the charismatic leadership of Pope Francis during the pandemic emergency; in order to better understand his effectiveness, we analyzed both the core issues and his multimodal body signals in the global TV event of the Universal Prayer with the *Urbi et Orbi* Blessing. The multimodal and discursive analyses of the homily enabled us to define the "humble" charisma of the Pope, which is based upon on authentic and informal presence, manifested emotional signals (and, in particular commotion) showing features of equity and familiarity. From a discursive point of view, the common and overarching affiliation is constructed through a multiple focus on the "we" pronoun, which is constructed through socio-epistemic rhetoric. The results show how this integrated methodological perspectives, which is multimodal and discursive, may offer meaningful pathways detection of effective and persuasive signals.

Keywords: COVID-19, charismatic leadership, humbleness, multimodal signals, discourse analysis, we-ness, Pope Francis

OPEN ACCESS

Edited by:
Willibald Ruch,
University of Zurich, Switzerland

Reviewed by:
Erik Cambria,
Nanyang Technological University, Singapore
Christian Cecconi,
Sapienza University of Rome, Italy

****Correspondence:***
Rosa Scardigno
rosa.scardigno@uniba.it

Specialty section:
This article was submitted to Personality and Social Psychology, a section of the journal Frontiers in Psychology

Received: *20 March 2021*
Accepted: *27 July 2021*
Published: *01 September 2021*

Citation:
Scardigno R, Papapicco C, Luccarelli V, Zagaria AE, Mininni G and D'Errico F (2021) The Humble Charisma of a White-Dressed Man in a Desert Place: Pope Francis' Communicative Style in the Covid-19 Pandemic. Front. Psychol. 12:683259. doi: 10.3389/fpsyg.2021.683259

INTRODUCTION

The coronavirus epidemic was declared "COVID-19 pandemic" by the World Health Organization in March 2020. As a serious global health emergency, the main effective countermeasure across the first year was "social distancing" undermining the sense of being part of a community. If we consider the COVID-19 pandemic as a kind of watershed in the modern history of humanity, the acronym "B.C." (before Christ), used in the Western world to indicate the whole period preceding the birth of Christ, could re-semanticized to designate the human condition "before COVID-19."

Among the several subjective resources and cultural anchorages, religious forms of life (Belzen, 2005) can support believers in facing this unprecedented scenario. As opposed to philosophical

and sociological claims about modernization and secularization, the essential role of religiosity in human life persists, as can be shown by the percentage of the world population still believing in God and by the strengthening role of religion within some societies (Bentzen, 2020). If in ordinary time, religions act as powerful "systems of meaning" (Park, 2010) addressing ultimate issues and concerns, in times of crisis human beings can witness a "turn" to religious contents and practices for comfort and explanations. As a matter of fact, almost all countries showed significant rises in prayer search; in March 2020, one of the most widespread kinds of prayer was the Coronavirus one, by means of which people asked for protection and strength and offered gratitude (Bentzen, 2020). A more general sign of the relation between the pandemic global fear and the religious "refuge" can be found in the considerable rise of a spiritual search (Maccioni, 2020).

The "searching" pathways can be empowered by the presence of charismatic leaders, since they concur to define and maintain social identity for their group (Hogg, 2001) and provide direction during times of uncertainty (Rast, 2015). In this perspective, the charismatic leadership of the Pope can play a crucial role in supporting prayers and in facing emergency with existential reasons and psychological resources. If we asked people to recollect the most significant and touching public images relating to COVID-19, probably most of them all over the world would talk about a lonely white-dressed man walking under the rain in a large empty square of the Eternal City. They would talk about Pope Francis celebrating the Universal Prayer with the *Urbi et Orbi* Blessing on March 27th. The enormous scope of this media event inspired social scholars and leaders with insights about the role of religious beliefs and practices, both in ordinary and extraordinary life events.

A recent study defined Pope Francis as a "digital leader" (Narbona, 2016), demonstrating the use that he made of his Twitter account for different areas, not only for the transmission of Christian values, but also for the proposal of specific actions and the fulfilling of more general educational tasks to both Catholic/Christian believers and people in general.

How is the leadership of the Pope characterized "offline" and how is it described in the COVID time? The present work aims to address some issues about communicative (both linguistic and bodily) features of the charismatic leader of the Christian Church during the pandemic emergency.

THE EFFECTIVE LEADERSHIP: CHARISMA OF MIND AND BODY

The socio-psychological literature assumes that charismatic leadership can be communicated by means of discursive practices (Seyranian and Bligh, 2008; Bligh and Kohles, 2009; de Vries et al., 2010) and non-verbal signals (Atkinson, 1984; Rosenberg and Hirschberg, 2009), pointing out that we can acknowledge a charisma of the mind and a charisma of the body (D'Errico et al., 2013).

As for the "charisma of mind," recent literature outlines alternative approaches to the more traditional "all-about-me" studies; in line with the "New Psychology of Leadership" (Haslam et al., 2011), an effective leadership concerns leaders and followers as joining the same membership and fulfilling the same needs. In the more general idea of leadership as a process rather than as an individual property, this group process is grounded in a shared social identity or "we-ness," emphasizing the psychological bond and the emotional leader–followers connection (Parry et al., 2019). This means that the effective leaders should encourage the collective sense of "who we are" and "what we are about" (Haslam and Reicher, 2016). Being both shared identity and shared reality at stake, the leader can gain influence by "shaping what we believe, what we value and how we should act" (Reicher et al., 2018, 129). In line with this model, a very recent review argues that instances of success and failure by different leaders during the pandemic are due to their ability to embody the shared interests of groups (Haslam et al., 2021).

The "charisma of the body" can be seen as a set of internal features of a person that, when manifested by some external displays, can influence people by inducing them to pursue some goals while feeling involvement and enthusiasm (D'Errico et al., 2013). Understanding charisma therefore means, on the one hand, to track the external displays in a multimodal perspective (Poggi, 2007) by analyzing words, prosody, voice, gesture, posture, face, gaze, body, and by decoding the expressed internal features.

The internal features expressed by the charismatic speaker can be potentially three: benevolence—caring for interests of others; competence—own expertise of an individual, planning capacity, and creativity; and dominance—the way to express own status position of an individual during the communicative event (D'Errico et al., 2013). In particular, the positioning of a speaker can be "vertical" and then communicate that she or he has more power/status than the others by means of several strategies (Poggi et al., 2011; D'Errico and Poggi, 2012); it can also communicate a horizontal/proximal stance toward the interlocutor by expressing on the contrary his/her own "humility" (D'Errico and Poggi, 2019). Although in the former case the speaker can raise the voice, interrupt, or overlap the interlocutor, in the latter case she or he can use longer pauses, lower intonation, and only occasionally incur in speech overlap, since she/he is awaiting his/her own turn to speak (D'Errico et al., 2019). They often look downward, with restrained gesture and tend to express negative emotions—especially sadness—while communicating their concern to the other (D'Errico, 2019). Generally speaking, humble stance is a "realistic" approach implying the awareness of being fallacious and imperfect. Despite humility being rather neglected in leadership literature (Collins, 2001), it is a fundamental "virtue of the heart"; more recently, it has been identified as an essential requirement of charismatic leadership, since it gives leaders a "charismatic touch" (Juurikkala, 2012). As a matter of this fact, in line with the Elicit–Channel (EC) model of charismatic leadership (Sy et al., 2018), leaders can both produce highly inspiring emotions from their followers and turn those emotions in actions that, if successful, result in positive outcomes, such as positive affect and trust.

This attitude implies several ones, such as (a) neglecting relevance to any external ornament like status symbols

(Essentiality); (b) proposing a feeling/communication of equality to others (Equality feature); and (c) neither feeling nor displaying any superiority (Non-Superiority feature).

THE RESEARCH

Aims and Research Questions

Within this framework, the overall goal of the present study is to investigate the charismatic leadership of Pope Francis through his multimodal communication. In particular, we tried to investigate his role of (religious) leader capable of harmonizing "contents" and "Gestalt," and to identify his communicative features that help audiences cope with the COVID-19 emergency period.

In this social scenario characterized by deep uncertainty and fear, we suppose that his charismatic leadership will be exercised by focusing on a core issue, that is a common higher-order belonging; despite his being the supreme spiritual head of Catholicism, feelings of brotherhood and equality can be widespread by claiming the uniqueness of God and the coincidence between "His people" and "Humanity."

This topic can be communicated through effective and persuasive communication acted by converging multimodal signals: as for the "charisma of mind," we expect that discursive practices will be oriented to emphasize the value of "brotherhood" as a shared horizon for the highlighted problems and for their solutions (Pope, 2020a); as for the "charisma of body," we hypothesize that his multimodal communication will emphasize the benevolence and the closeness to sufferings of others (empathy). A charismatic mind and body will be characterized by a humble communicative style.

Procedures and Data

The present qualitative study is based on the multimodal analysis of a unique event—that is the Universal Prayer with the *Urbi et Orbi* Blessing on March 27th—which was transmitted live all over the world. Although the whole liturgy lasted for about 1 h and 24 min, the homily took about 16 min. This part of the liturgy was transcribed and analyzed through a multimodal approach focusing on body feature annotation and on paper-and-pencil discourse analysis.

Multimodal Communication Analysis

In multimodal communication analysis, the facial expressions, gestures, and posture of the Pope were taken into consideration. As for facial expressions, both the Facial Action Coding System (FACS) coding model (Ekman, 1993) and Action Units (AUs) detection were also used. The selected video was included in ELAN (Max Planck Institute for Psycholinguistics, Nijmegen; Wittenburg et al., 2006), that is, an assisted annotation software tool in which categories (called "tier") were created corresponding to each non-verbal communication signal.

After the annotation process, the following decoding step was the attribution of meaning to each non-verbal communication signal. Therefore, multimodality derives from the set of meanings of both verbal and non-verbal signals.

Hence, the result of this meticulous encoding and decoding procedure is not only composed of non-verbal signals with their related meanings according to humility features—that is familiarity, empathy, non-superiority, informality, equality, and essentiality (D'Errico and Poggi, 2019)—but also of the exact timing and signal duration.

Discourse Analysis

The homily of the Pope was analyzed by the means of diatextual analysis (Manuti et al., 2017), that is a type of critical discourse analysis focusing on the reciprocal co-construction of texts and contexts emerging from each discourse (Scardigno et al., 2019). In line with the overall aim of this work, we selected some specific diatextual markers mainly having to do with the argumentative-rhetoric construction of the homily. More specifically, we drew on "social-epistemic rhetoric" (Berlin, 1993), that is a discourse construction that permits a top-down reading of texts by grasping sense perspectives valid for specific positioning groups.

An overview of the data was obtained through a semiotic square (Greimas and Courtés, 1979): it is a semiotic device enabling researchers to show the oppositional logic concerning every concept. In particular, the discourse enunciative positions can be organized upon three axes: opposition (white vs. black), contradiction (white vs. non-white), and subopposition (non-white vs. non-black). As such, the semiotic square is the concrete sign of the dialog between discursive data and interpretative options acted by social researchers.

Main Results Discourse Analysis

Multimodal Analysis

In the analyzed communicative event, the multimodal communication of Pope Francis revealed high coherence among its diverse signals: no inconsistencies among facial expressions, posture, and gestures were found. This could be interpreted as an authentic attitude and spontaneous emotional pattern, despite its being an "official" celebration.

An in-depth analysis across the minutes of the matching among signals and meanings enabled us to outline two essential features: (a) commotion as the main noticeable emotion and (b) the emphasis upon the verbal, through facial expressions and gestures. The average duration of the detected multimodal signals in this ceremony is 224 thousandths of a second out of 28 min of analyzed video. In the context of an official ceremony, the posture of the Pope is "open," his shoulders turned to the audience that is ideally present, the ritual gestures are slow but as precise as the speech rate, he speaks in a low-pitched voice by clearly pronouncing the words. At the beginning of the homily, the Pope moves toward the altar, goes up and, with a trembling voice, and raises his eyes to heaven in order to give full sense to prayers and to ask for protection ("Through Christ our Lord") (**Figure 1**).

Such recurrent emotional signals as commotion follow sentences like "We found ourselves frightened and lost" (min. 10.58) expressing an emotional closeness to the suffering of other people. **Figure 2** showed a significant example of commotion AUs detecting, such as AU15 + AU23 (Kimura et al., 2019): muscle contraction mainly involves the lips of the Pope. In particular, the outer part lowers and creates some sort of arc.

FIGURE 1 | Pope Francis (minute 6.18).

FIGURE 2 | Pope Francis commotion by face (minute 10).

FIGURE 3 | Emphatic facial expression (minute 15).

If the lip contraptions were decontextualized from speech and other signs, the identified AUs would be sadness. Nonetheless, in this case the movement of the head bends forward, not only to read the script, with a movement of 0.56 thousandths of a second. In addition, the corresponding verbal expression refers to the difficult period of the pandemic. As a consequence, the set of these multimodal signals can be holistically interpreted as "commotion": the emotion of "movingness" is by definition included neither in joy nor in sadness; it rather introduces additional features found in highly selected episodes of sadness and joy (Kuehnast et al., 2014). An event of sadness arousing emotion is precisely the period of pandemic which aggravated uncertainty and fear during its first wave.

The moved and authentic emotion expression acted by the Pope represents a starting point to define him as a humble leader (Owens et al., 2019). The non-manipulation of emotionality, *albeit* in a formal ceremony, provides the audience with the image of someone feeling first man and then the Pope: in the guise of a man, he feels like the others and, in those of the Pope, he empathizes the mankind (*Equality feature*).

Features of "humility" are also outlined through the emphasis upon some discursive excerpts. In particular, batonic gestures (Navarretta, 2018)—for example movements of arms from the top to the bottom—are recurrent in the combination with expressions such as "We are." The combined use of the first plural person and of the emphasizing batonic gestures (with one hand or both) empowers the feelings of *equality* and *familiarity*.

As for facial expressions, eyebrows contraction, especially in the inner part (AU4), can be observed (**Figure 3**). As already shown in **Figure 1**, this facial sign aims to strengthen and complete what is said (in this case "Always").

In the central and final parts of the homily, the unit AU4 was found more frequently. Analyzing the verbal part being expressed synchronically with AU4 can help us to better understand the several meanings associated to this facial signal (**Table 1**); it can create a feeling of familiarity by emphasizing the sense of closeness (as in the case of "To enhance," timing 00:25:08.709-00:25:08.929); otherwise, it can convey a mixture of empathy and slight anger by proposing a kind of "admonition" (as in the case of "The others say about us," timing 00:20:09.205-00:20:09.425). Even in its variety, AU4 concurs to create the atmosphere of careful closeness, which strengthens the image of the Pope as a humble leader.

Discourse Analysis

In the particular scenario in which the mediatic event takes place, and in line with the multimodal communication of the Pope, discourse analysis reveals that his homily focuses on a really pathemic argumentation, where humility is mainly emphasized through equality attitudes.

Through a continuous contamination between the religious and secular domains, between "religionese" (Scardigno and Mininni, 2020)—that is the language of religion—and common speech, positive and negative connotations, "up" and "down" attitudes, what is at stake in this homily is a kind of "egalitarianism" overcoming religious borders and belongings, including all human beings. As a matter of fact, the Pope almost bans the "I" subject pronoun in favor of the first person plural, so that the whole discourse is dotted with "we": subjects and verbs,

adjectives, and pronouns testify to his commitment in the several experiences and connotations of human belonging.

In order to have both a synthetic and analytic vision of his "deep" discourse, a semiotic square was constructed (**Figure 4**). In line with the multiple dialectic and oppositive dimensions, we found the following oppositions:

(a) The basic opposition between "weak" and "strong" relations. Feelings, attitudes, and relations of human beings can be set in this basic opposition, concerning the nature of involvement in what is strictly "human." Weak relations have to do with superficial attitudes and links, whereas strong relations with more careful and mindful ones.
(b) The consequent contradictions between "weak"/"non-weak" and "strong"/"non-strong" relations, involving different nuances of being involved and explaining the modulated "deepness" in constructing relations.
(c) The resulting subopposition between "non-weak" and "non-strong" relations, concerning the alternative quality of the nuanced activated relations.

The dilemmatic oppositions emerging from the semiotic square represent the axis upon which four types of "we" can be set:

(a) "weak": when the Pope talks about this position and relations, he makes use of verbs and adjectives concerning human selfishness (e.g., "greedy for gain," "we have not heard the cry of the poor") and superficiality (e.g., "we let ourselves be absorbed by objects and stunned by hurry"), which are typical of the "Pre-COVID" time.
(b) "non-strong": in this positioning, relations are founded on limited attitudes and exhibited through the unreadiness of humanity (e.g., "we found ourselves frightened and lost") in facing with the global and social emergency during "COVID-time."
(c) "strong": this kind of "we" is elaborated through the discursive act of exhortation. As a matter of fact, in order to occupy this position, which highlights worthiness and emphasizes "solid relations", the Pope aims for a "faithful" *we*, based on serious religious commitment:

"let's invite Jesus in the boats of our life. Let's give Him our fears, so that He can overcome them".

(d) "non-weak": as in type (c), this position can be set through the adhesion to more philanthropic proposals, which can transform "weak" relations in "non-weak" ones, aspiring to a "possible" *we*:

"[. . .] that common belonging from which we cannot escape, that is belonging as brothers."

In such a multifaceted articulation, the Pope proposes a "religious" opposition between weak/sinful and strong/faithful *we* thus implying the "human" opposition between non-strong/unready and non-weak/possible *we*. In addition, the axis of weak/non-strong and strong/non-weak relations can be summarized and represented through two social-epistemic rhetorics, "fallibility" and "worthiness," acting as frames for the opposite attitudes of common limitation, that is, faultiness, inadequacy, and imperfection of human beings, and full expression of human potential, based on the aspiration to goodness and rightness.

As for the pathemic connotation of discourse, the following additional discursive markers were found:

(a) *Metaphors*. In particular, three main kinds were used: (1) common speech metaphors, for example, "we went full speed," referring to human and common habits; (2) religious metaphors, for example, "to embrace His Cross," mainly deriving from the Christian symbolism; and (3) archetypal metaphors, for example, the "evening" representing the end of the day and life.
(b) Other *figures of speech*, such as climax, repetition, and the tripartite list:

"We have an anchor: in His Cross we were saved. We have a helm: though His Cross we were redeemed. We have a hope: in His Cross we were restored and embraced".

These figures both creates an emotional atmosphere and improve the public involvement.

(c) *Polarized adjectives*, for example, "all of us fragile and disoriented but, at the same time, important and necessary." The discursive politization acts as an additional emotional charge.

Both rhetoric figures and affective markers work in order to empower the empathic activation by increasing the perception of a strong relationship and of a humble leader.

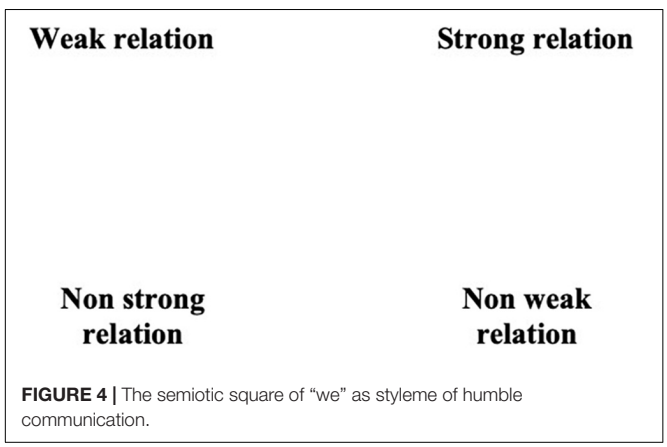


FIGURE 4 | The semiotic square of "we" as styleme of humble communication.

TABLE 1 | Recurrence of AU4 emphasis expression and corresponding speech.

Timing	Speech
00:15:33.410-00:15:33.630	Pause
00:20:09.205-00:20:09.425	"The others say about us"
00:24:56.691-00:24:57.041	"Always"
00:25:08.709-00:25:08.929	"To enhance"
00:27:53.739-00:27:53.909	"To open spaces"

DISCUSSION

The present qualitative analysis represented an attempt to investigate how the Pope proposed his undoubted charisma through a sensible mixture of discursive practices and multimodal signals making him a coherent and deep "living human document" (Ganzevoort, 1993) for an audience virtually coinciding with humankind, during an exceptional media event in the full pandemic emergency.

Since his first public appearance as a Pope, Jorge Mario Bergoglio characterized his communicative style through the "rhetoric of the unexpected." Indeed, "the frame 'first greeting of the Pope' is reinvented—distances are shortened and hearts touched. From that day on, Pope Francis has often displayed unexpected behaviors" (Caffi, 2015, 23).

Since the "unexpected" magisterium and the pastoral action of Pope Bergoglio was also marked by the words "meeting" and "people" (Pope, 2020b), we supposed that in the selected event he would try to offer empowered pathways to highlight these concepts. The multimodal and discursive analysis of the homily enabled us to construct this type of frame; on the one hand, the work carried out with the ELAN-assisted annotation software tool revealed an authentic and informal presence, with manifested emotional signals, even in a formal celebration. The eyes, face, postures, and batonic gestures analyzed in their contexts are signals revealing both feelings of commotion, and equal, familiar, and emphatic stance. On the other hand, discourse analysis emphasized the performative strength of discursive practices in line with a pathemic argumentation based on metaphors and other rhetoric devices.

Common and overarching belonging is implemented through a multiple focus on "we," which is constructed starting from the basic opposition between "weak" and "strong" relations. Looking at the four positions emerging from the semiotic square, a temporal axis can be outlined: while (a) is almost referred to the past and to the general attitudes/behaviors before the COVID-19 pandemic, (b) involves a particular actual time, that is the COVID-19 chronotope; both (c) and (d) imply a near-future time, that is welcome in line with the exhorted/proposed attitudes/behaviors. Nonetheless, the positions (c) and (d) are slightly different: the faithful *we* is exhorted on the basis of the confirmation of religious feelings and of the explicit presence of God in one's own and common life; the possible *we* is founded on the (new) opportunity to construct worth and non-weak relations, that is on civic engagement and human care.

Taking into account the "twelve lessons" of the effective leadership associated with the crisis management proposed by Haslam et al. (2021), this work revealed the presence of several elements, such as the construction of a shared identity, the proposal of inclusive definition of in-groups, an empathic attitude, and so on. In a multifaceted and shared construction of human attitudes and behaviors, the Pope takes on his shoulders the responsibility of what is wrong and the opportunity of what can get better, thus defining at the same time not only "who we are" and "what we are about" (Haslam and Reicher, 2016), but first and foremost "what we can be" (the possible *we*) and "what we should be" (the faithful *we*).

Nonetheless, the multimodal communication of the Pope is unique in proposing a coherent model in the direction of humbleness. Confirming the unicity of "COVID-19 pandemic" as a kind of watershed in the modern history of humanity, both multimodal and discourse analyses of the homily of the "Urbi et Orbi" media event enabled us to consider the features of Pope Francis as a "humble leader": his emphatic and moved attitude enhanced commitment; his charisma was specifically based on careful closeness and fraternal humanism and on an egalitarian contamination between religious and non-religious lexicon and belonging.

CONCLUDING REMARKS

Since the charismatic leadership of Pope Francis is unquestionable both in the religious domain and in public opinion, the general aim of this work was to analyze how communication signals can play a significant role in His charisma, especially in such times of crisis and uncertainty as the COVID-19 pandemic. Social psychology of communication can offer a meaningful perspective for a conjoint focus on multimodal signals and discursive practices through the analysis of effective communicative performances. In order to better appreciate the several features of the charismatic leadership of Pope Francis, more extensive analysis on other communicative events should be carried out. Nonetheless, the proposed combined study can represent a novel methodological attempt to deeply analyze the associated meaning alongside linguistic and bodily signals. Specifically, the methodology used, which consists of multimodal and discursive analyses, can fall within the field of Social Signal Processing (Vinciarelli et al., 2011) and of Sentic Computing (Susanto et al., 2021). This effort has been applied also within the field of automatic signal detection, which recently have crossed both linguistic levels, as in the case of Sentiment Multimodal Analysis (Tiwari et al., 2020), and other body channel detection (e.g., facial or vocal expressions and body postures) (Soleymani et al., 2017). More in general, the recent remarkable development of the Affective Computing and the introduction of deep learning increased the development of more sophisticated systems (Wadawadagi and Pagi, 2020) aimed at predicting and automatically detecting emotional patterns starting from multimodal input (Morency et al., 2011; Poria et al., 2016; Chaturvedi et al., 2019; Yadav and Vishwakarma, 2019), with a real-time output (Cambria et al., 2013; Tran and Cambria, 2018). In this perspective, future studies will try to apply automatic systems to wider textual and visual corpora, promoting therefore interdisciplinary approaches.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

AUTHOR CONTRIBUTIONS

RS as the first author, was responsible for the coordination of the whole project; in addition, she wrote a part of the introduction and of Sections "The Effective Leadership: Charisma of Mind and Body" and "The Research"; and she conducted and wrote the discourse analysis section and discussed its findings. CP conducted the multimodal communication analysis and reported results. VL wrote a part of Section "The Effective Leadership: Charisma of Mind and Body" and supported literature review; in addition, she took part in the discourse analysis phase. AZ took part in the discourse analysis process and wrote the concluding section of the manuscript. GM as a senior author, supervised the whole work; in addition, he wrote a part of the introduction and of Sections "The Effective Leadership: Charisma of Mind and Body" and "The Research". FD'E supervised the theoretical, methodological, and analytical sections devoted to multimodal communication; in addition, she wrote part of Sections "The Effective Leadership: Charisma of Mind and Body" and "The Research". All authors contributed to the article and approved the submitted version.

FUNDING

Funds were provided by the Administrative Board of the University of Bari.

ACKNOWLEDGMENTS

The authors would like to thank the colleagues at the Department of Educational Science, Psychology, Communication of the University of Bari for the fruitful feedback provided to this research proposal.

REFERENCES

Atkinson, M. (1984). *Our Masters' Voices: The Language and Body Language of Politics*. London: Psychology Press.

Belzen, J. A. (2005). A way out of crisis? From Völkerpsychologie to cultural psychology of religion. *Theory Psychol.* 15, 812–838. doi: 10.1177/0959354305059334

Bentzen, J. (2020). *In Crisis, We Pray: Religiosity and the COVID-19 Pandemic. CEPR Discussion Papers 14824*. London: CEPR.

Berlin, J. A. (1993). "Post-structuralism, semiotics, and social-epistemic rhetoric: converging agendas," in *Defining the New Rhetoric*, eds T. Enos and S. Brown (Newbury Park, CA: Sage), 137–176. doi: 10.1057/978-1-137-59837-0_7

Bligh, M. C., and Kohles, J. C. (2009). "Charismatic Leadership," in *Encyclopedia of Group Processes and Intergroup Relations*, eds J. M. Levine and M. A. Hogg (London: Sage).

Caffi, C. (2015). (Un)expected behaviour: some general issues and a papal example. *J. Pragmat.* 86, 19–24. doi: 10.1016/j.pragma.2015.06.008

Cambria, E., Howard, N., Hsu, J., and Hussain, A. (2013). "Sentic blending: scalable multimodal fusion for the continuous interpretation of semantics and sentics," in *Proceedings of the 2013 IEEE Symposium on Computational Intelligence for Human-Like Intelligence (CIHLI)* (Piscataway, NJ: IEEE), 108–117.

Chaturvedi, I., Satapathy, R., Cavallari, S., and Cambria, E. (2019). Fuzzy commonsense reasoning for multimodal sentiment analysis. *Pattern Recognit. Lett.* 125, 264–270. doi: 10.1016/j.patrec.2019.04.024

Collins, J. (2001). Level 5 leadership: the triumph of humility and fierce resolve. *Harv. Bus. Rev.* 79, 66–76.

de Vries, R. E., Bakker-Pieper, A., and Oostenveld, W. (2010). Leadership = communication? The relations of leaders' communication styles with leadership styles, knowledge sharing and leadership outcomes. *J. Bus. Psychol.* 25, 367–380. doi: 10.1007/s10869-009-9140-2

D'Errico, F. (2019). 'Too humble and sad': the effect of humility and emotional display when a politician talks about a moral issue. *Soc. Sci. Inform.* 58, 660–680. doi: 10.1177/0539018419893564

D'Errico, F., and Poggi, I. (2012). Blame the opponent! Effects of multimodal discrediting moves in public debates. *Cogn. Comput.* 4, 460–476. doi: 10.1007/s12559-012-9175-y

D'Errico, F., Niebuhr, O., and Poggi, I. (2019). "Humble voices in political communication: a speech analysis across two cultures," in *Proceedings of the International Conference on Computational Science and Its Applications* (Cham: Springer), 361–374. doi: 10.1007/978-3-030-24296-1_30

D'Errico, F., and Poggi, I. (2019). Tracking a leader's humility and its emotions from body, face and voice. *Web Intell.* 17, 63–74. doi: 10.3233/web-190401

D'Errico, F., Signorello, R., Demolin, D., and Poggi, I. (2013). "The perception of charisma from voice: a cross-cultural study," in *Proceedings of the 2013 Humaine Association Conference on Affective Computing and Intelligent Interaction* (Piscataway, NJ: IEEE), 552–557.

Ekman, P. (1993). Facial expression and emotion. *Am. Psychol.* 48:384.

Ganzevoort, R. R. (1993). Investigating life-stories. personal narrative in pastoral psychology. *J. Psychol. Theol.* 21, 277–287. doi: 10.1177/009164719302100401

Greimas, A. J. e, and Courtés, J. (1979). *Sémiotique. Dictionnaire Raisonné de la Théorie du Langage*. Paris: Hachette.

Haslam, S. A., and Reicher, S. D. (2016). Rethinking the psychology of leadership: from personal identity to social identity. *Daedalus* 145, 21–34. doi: 10.1162/daed_a_00394

Haslam, S. A., Reicher, S. D., and Platow, M. J. (2011). *The New Psychology of Leadership–Identity, Influence and Power*. East Sussex: Psychology Press.

Haslam, S. A., Steffens, N. K., Reicher, S. D., and Bentley, S. V. (2021). Identity leadership in a crisis: a 5R framework for learning from responses to COVID-19. *Soc. Issues Policy Rev.* 15, 35–83. doi: 10.1111/sipr.12075

Hogg, M. A. (2001). A social identity theory of leadership. *Pers. Soc. Psychol. Rev.* 5, 184–200. doi: 10.1207/s15327957pspr0503_1

Juurikkala, O. (2012). Likeness to the divinity? Virtues and charismatic leadership. *Electron. J. Bus. Ethics Organ. Stud.* 17, 4–15.

Kimura, K., Haramizu, S., Sanada, K., and Oshida, A. (2019). Emotional state of being moved elicited by films: a comparison with several positive emotions. *Front. Psychol.* 10:1935. doi: 10.3389/fpsyg.2019.01935

Kuehnast, M., Wagner, V., Wassiliwizky, E., Jacobsen, T., and Menninghaus, W. (2014). Being moved: linguistic representation and conceptual structure. *Front. Psychol.* 5:1242. doi: 10.3389/fpsyg.2014.01242

Maccioni, R. (2020). *Il Sociologo Garelli: 'al Tempo del Covid Cresce il Bisogno di Dio'*. Available online at: https://www.avvenire.it/chiesa/pagine/cos-cresce-il-bisogno-di-dio (accessed April 19, 2020).

Manuti, A., Scardigno, R., and Mininni, G. (2017). Diatexts at work: the diatextual approach as a psycho-cultural path of critical discourse analysis in the organizational context. *Qual. Res. J.* 17, 319–334. doi: 10.1108/qrj-07-2016-0044

Morency, L. P., Mihalcea, R., and Doshi, P. (2011). "Towards multimodal sentiment analysis: harvesting opinions from the web," in *Proceedings of the 13th International Conference on Multimodal Interfaces*, Alicante, 169–176.

Narbona, J. (2016). Digital leadership, twitter and Pope Francis. *Church Commun. Cult.* 1, 90–109. doi: 10.1080/23753234.2016.1181307

Navarretta, C. (2018). "The automatic annotation of the semiotic type of hand gestures in Obama's humorous speeches," in *Proceedings of the 11th International Conference on Language Resources and Evaluation* (LREC 2018), Paris.

Owens, B. P., Yam, K. C., Bednar, J. S., Mao, J., and Hart, D. W. (2019). The impact of leader moral humility on follower moral self-efficacy and behavior. *J. Appl. Psychol.* 104:146. doi: 10.1037/apl0000353

Park, C. L. (2010). Making sense of the meaning literature: an integrative review of meaning making and its effects on adjustment to stressful life events. *Psychol. Bull.* 136, 257–301. doi: 10.1037/a0018301

Parry, K., Cohen, M., Bhattacharya, S., North-Samardzic, A., and Edwards, G. (2019). Charismatic leadership: beyond love and hate and toward a sense of belonging? *J. Manag. Organ.* 25, 398–413. doi: 10.1017/jmo.2019.28

Poggi, I. (2007). *Mind, Hands, Face and Body: A Goal and Belief View of Multimodal Communication*. Berlin: Weidler.

Poggi, I., D'Errico, F., and Vincze, L. (2011). "Discrediting moves in political debate," in *Proceedings of the Second International Workshop on User Models for Motivational Systems: The Affective and the Rational Routes to Persuasion*, 84–99.

Pope, F. (2020a). *Fratelli tutti. Enciclica sulla fraternità e l'amicizia sociale*. Roma: Libreria Editrice Vaticana.

Pope, F. (2020b). *Ritorniamo a Sognare. La Strada Verso un Future Migliore*. Milano: Piemme.

Poria, S., Cambria, E., Howard, N., Huang, G. B., and Hussain, A. (2016). Fusing audio, visual and textual clues for sentiment analysis from multimodal content. *Neurocomputing* 174, 50–59. doi: 10.1016/j.neucom.2015.01.095

Rast, D. E. (2015). Leadership in times of uncertainty: recent findings, debates, and potential future research directions. *Soc. Personal. Psychol. Compass* 9, 133–145. doi: 10.1111/spc3.12163

Reicher, S. D., Haslam, S. A., and Platow, M. J. (2018). Shared social identity in leadership. *Curr. Opin. Psychol.* 23, 129–133.

Rosenberg, A., and Hirschberg, J. (2009). Charisma perception from text and speech. *Speech Commun.* 51, 640–655. doi: 10.1016/j.specom.2008.11.001

Scardigno, R., and Mininni, G. (2020). An authentic feeling? Religious experience through Q&A websites. *Arch. Psychol. Relig.* 42, 211–231. doi: 10.1177/0084672420917451

Scardigno, R., Pastore, S., and Mininni, G. (2019). The mixed family as a discursive melting pot of cultures [La famiglia mista come crogiolo discorsivo di culture]. *Psicol. Soc.* 14, 73–98.

Seyranian, V., and Bligh, M. C. (2008). Presidential charismatic leadership: exploring the rhetoric of social change. *Leadersh. Q.* 19, 54–76.

Soleymani, M., Garcia, D., Jou, B., Schuller, B., Chang, S. F., and Pantic, M. (2017). A survey of multimodal sentiment analysis. *Image Vis. Comput.* 65, 3–14. doi: 10.1016/j.imavis.2017.08.003

Susanto, Y., Cambria, E., Ng, B. C., and Hussain, A. (2021). Ten years of sentic computing. *Cognit. Comput.* 1–19.

Sy, T., Horton, C., and Riggio, R. (2018). Charismatic leadership: eliciting and channeling follower emotions. *Leadersh. Q.* 29, 58–69. doi: 10.1016/j.leaqua.2017.12.008

Tiwari, S., Verma, A., Garg, P., and Bansal, D. (2020). "Social media sentiment analysis on twitter datasets," in *Proceedings of the 2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS)* (Piscataway, NJ: IEEE), 925–927.

Tran, H. N., and Cambria, E. (2018). Ensemble application of ELM and GPU for real-time multimodal sentiment analysis. *Memet. Comput.* 10, 3–13. doi: 10.1007/s12293-017-0228-3

Vinciarelli, A., Pantic, M., Heylen, D., Pelachaud, C., Poggi, I., D'Errico, F., et al. (2011). Bridging the gap between social animal and unsocial machine: a survey of social signal processing. *IEEE Trans. Affect. Comput.* 3, 69–87. doi: 10.1109/t-affc.2011.27

Wadawadagi, R., and Pagi, V. (2020). Sentiment analysis with deep neural networks: comparative study and performance assessment. *Artif. Intell. Rev.* 53, 6155–6195. doi: 10.1007/s10462-020-09845-2

Wittenburg, P., Brugman, H., Russel, A., Klassmann, A., and Sloetjes, H. (2006). "ELAN: a professional framework for multimodality research," in *Proceedings of the 5th International Conference on Language Resources and Evaluation (LREC 2006)*, Genoa, 1556–1559.

Yadav, A., and Vishwakarma, D. K. (2019). Sentiment analysis using deep learning architectures: a review. *Artif. Intell. Rev.* 53, 4335–4385. doi: 10.1007/s10462-019-09794-5

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Scardigno, Papapicco, Luccarelli, Zagaria, Mininni and D'Errico. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Psychology

ORIGINAL RESEARCH
published: 01 September 2021
doi: 10.3389/fpsyg.2021.695468

Embedded Philanthropic CSR in Digital China: Unified View of Prosocial and Pro-environmental Practices

Qing Ye[1,2,3], Zain Rafique[4], Rongting Zhou[1,3], Fahad Asmi[1,3], Muhammad Azfar Anwar[5,6]* and Ahmad Nabeel Siddiquei[7]*

[1] Department of Science and Technology of Communication, University of Science and Technology of China, Hefei, China, [2] College of Information Engineering, Fuyang Normal University, Fuyang, China, [3] Key Laboratory of Immersive Media Technology (Anhui Xinhua Media Co., Ltd.), Ministry of Culture and Tourism, Hefei, China, [4] Department of Governance and Public Policy, National University of Modern Languages, Islamabad, Pakistan, [5] Research Institute of Business Analytics and Supply Chain Management, College of Management, Shenzhen University, Shenzhen, China, [6] Department of Management Sciences, COMSATS University Islamabad, Vehari, Pakistan, [7] Bond Business School, Bond University, Gold Coast, QLD, Australia

In recent decades, China has transformed from a conventional society into a digitally competitive nation. From an economic perspective, embedded corporate social responsibility (CSR) is gaining a new height where gamified charity is a trendy approach. By adopting the norm activation model from the point of view of the stimulus–organism–response framework, this research theoretically conceptualized the role of the mobile application environment (including telepresence, functional transparency, and accessibility) to map the cognition and philanthropic behavioral intentions of consumers in the gamified setting. The quantified survey comprised 669 respondents. The findings highlighted the critical role of functional transparency and telepresence of a mobile application in driving consumers' warm glow and ascribed responsibility. The research underlined the presence of the unique DNA of Internet Plus Charity (Public Benefits) for prosocial and pro-environmental purposes in China under the umbrella of philanthropic CSR.

Keywords: CSR, gamified charity, S-O-R framework, functional transparency, telepresence, warm glow

OPEN ACCESS

Edited by:
Francesca D'Errico, University of Bari Aldo Moro, Italy

Reviewed by:
Laurette Dubé, McGill University, Canada
Concetta Papapicco, University of Bari Aldo Moro, Italy

****Correspondence:***
Fahad Asmi
fasmie@ustc.edu.cn
Muhammad Azfar Anwar
azfaranwar@ciitvehari.edu.pk

Specialty section:
This article was submitted to Human-Media Interaction, a section of the journal Frontiers in Psychology

Received: *15 April 2021*
Accepted: *28 June 2021*
Published: *01 September 2021*

Citation:
Ye Q, Rafique Z, Zhou R, Asmi F, Anwar MA and Siddiquei AN (2021) Embedded Philanthropic CSR in Digital China: Unified View of Prosocial and Pro-environmental Practices. Front. Psychol. 12:695468. doi: 10.3389/fpsyg.2021.695468

INTRODUCTION

Corporate social responsibility (CSR) provides a broadly coupled scope of organizations' responsibilities on the ethical, social, and organizational fronts (Bian et al., 2021). It reflects that shareholder economic benefits must come up in a sustainable manner (Matten and Moon, 2008). Kao et al. (2018) mentioned that CSR practices in mainland China are primarily influenced by present external factors (i.e., being part of the global supply chain) and internal influencers (i.e., institutional code of conduct). Companies in China strive to gain market share while keeping their social and ecological concerns balanced and well addressed. Interestingly, the reputation of non-state-owned enterprises is highly influenced by the degree of social and environmental standards the firms have achieved (Barnea and Rubin, 2010; Zhang et al., 2015). Yu et al. (2017) also indicated that CSR positively promotes a firm's competitive advantage among competitors. The Chinese government has legalized CSR into a law framework, indicating that CSR is widely accepted in China while in the era of Industry 4.0 (Yu et al., 2017; Kao et al., 2018). In the context of corporations, CSR practices have been recognized as a prominent approach to improve

competitiveness; therefore, companies have increasingly chosen to embed social responsibility into their operations and management (Xiao and Yang, 2020) in which the new practice paradigm of corporate philanthropy practice is formed on the basis of the Internet (Xiao et al., 2020).

In mainland China, the concept of philanthropy is evolving while having unique Chinese characteristics, which broadly overlap prosocial and pro-environmental behaviors. Therefore, the combination of charity and public benefits can be interchangeably used for prosocial and pro-environmental issues and concerns. Their core value (service) generally addresses socioecological issues, that is, poverty alleviation, disaster relief, education, ecological and cultural conservation, and science and technology development. Furthermore, information and communications technology (ICT) can be regarded as a catalyst to increase philanthropic practices in society.

The tracking of the fusion of "ICT" and "philanthropy" comprises three stages. First, the Internet cooperates with philanthropy as a communication tool. Second, IT offers a communication function to philanthropy in which IT remains a subsidiary of philanthropy but integrates gradually. Third, the Internet and philanthropy have fully merged as an organic whole called "*Internet Plus Charity (Public Benefits)*" in China. The characteristics of the Internet, such as data, intelligence, cloud, open, interactive, and other IT-driven innovative concepts, have been used to extract the realized potential for philanthropic reforms. As a result, several new charitable participation (public benefit) modes have appeared in the electronic world. Another exciting change is that new models combine virtual and real practices to show the impact of the real world.

In the context of *Internet Plus Charity (Public Benefits),* the public has various engaging approaches to charity that have features of diversification, contextualization, and gamification, such as donation from reading, walking, a daily low-carbon lifestyle, and even usage or consumption of daily application programs. From the point of view of corporations, it is an emerging paradigm that embeds corporate philanthropy into product and operation; for example, "Ant Forest" by Alipay and "Fulfilling *Dream Elf*" by Toutiao. Such an emerging paradigm is called corporate philanthropic behavior, but also belongs to cause-related marketing. Meanwhile, it offers a convenient way for the public to join charity (public benefits). The current paper coins the term "gamified charity" to highlight the strategic use of games to drive and trigger charitable concerns (specifically, social benefits). That is, the ICT has minimized the effort required by potential initiators/contributors by providing electronic platforms to transform their intentions in a real-world manner. Research examines the potential explanatory power of accessibility (*ACC*), telepresence (*TEL*), and functional transparency (*TRA*).

In particular, the authors take the contextual constructs defined on the basis of application-environment properties to drive the philanthropic behaviors of potential consumers. That is, there is an enormous amount of literature existing on the philanthropic behavioral mapping established by the norm activation model (NAM). However, none of the existing literature highlights the role of virtual application attributes in predicting the philanthropic behaviors of potential consumers. Thus, the present study takes a unique position as the research objective to underline the significance of the attributes of the virtual application environment. *ACC*, *TEL*, and *TRA* while explaining the philanthropic intentions of consumers. Apart from this unique essence of the research, it further expands the mediating role of motivation-driven (warm glow, *WG*) and consequence-driven (ascribed responsibility, *AR*) cognitive aspects as determinants while exploring the relatedness between the attributes of the application environment and consumers' philanthropic intentions. In the recent literature pool, social and personal norms have been used as exogenous factors to define responsible behavior. To further expand the theoretical implication of the current research, the moderating role of the normative environment (NS) is investigated while determining the endogenous factor (philanthropic behavioral intentions) in mainland China.

CORPORATE PHILANTHROPY AND GAMIFIED CHARITY

Corporate social responsibility is a multidimensional construct that includes a range of corporate behaviors aimed at fulfilling the expectations of different stakeholders (Farop et al., 2016; Jia, 2020). In recent times, CSR has been examined in several businesses and economic environments, that is, human resource management (Kao et al., 2018), marketing differentiation (Yu et al., 2017), and portfolio planning and management (Geerts and Dooms, 2020). Furthermore, some studies have emphasized that internal CSR promotes employees' engagement in voluntary eco-concerned initiatives (Su and Swanson, 2019; Alsuwaidi et al., 2021). Nan and Heo (2013) indicated that cause-related marketing elicits a favorable consumer attitude toward corporate identity. The adoption of ICT is an effective method to effectively communicate CSR to stakeholders/potential participators (Du et al., 2010).

In general, the concept proposed by Carroll (1979) can define CSR, which includes four dimensions: economic, legal, ethical, and philanthropic responsibilities. Philanthropic responsibilities involve actions that connect with society's expectation (participating in society's welfare) (Carroll, 1991). As a representative behavior of CSR, corporate philanthropy can transmit the positive social values of an enterprise to internal and external stakeholders, thus improving the corporate image to form a competitive advantage (Zhou et al., 2019). In China, the corporate philanthropy development model has changed from "government-led and enterprise participation" to "Enterprise building platform and multisocial participation" (Xiao et al., 2020). Meanwhile, ICT-driven corporate philanthropy has gained popularity, especially among young people. Therefore, an increasing number of companies have launched gamified charity, which integrates gamified elements with charity to transform its potential value into the realized ones in a convenient manner. Specifically, interactive game design can be a crucial factor in attracting extensive attention and broad participation from the public.

Gamification means implementing game design elements in non-game contexts (Deterding et al., 2011). Gamification can stimulate the willingness of users to participate in any given task or activity (Huotari and Hamari, 2012). That is, gamification aims to enhance assigned charges or actions in an exciting manner (Koivisto and Hamari, 2014). The operating principle of gamification is to employ, engage, and reward users who produce cognitive, persuasive, and psychological effects (Koivisto and Hamari, 2019). Typical gamification forms include points, leaderboards, achievements, feedback, clear goals, and narratives (Hamari and Koivisto, 2015), which can be described as creating social competition and incentivizing behavior through badge and reward systems (Hanus and Fox, 2015).

In recent decades, gamification has been applied in various fields, such as business, marketing, e-Learning, digital healthcare, and individually and socially sustainable behaviors (Koivisto and Hamari, 2014; Robson et al., 2015; Sardi et al., 2017). Meanwhile, CSR can be captioned as a necessity to differentiate in the highly competitive environment and further lead promotional strategies (Chen and Yang, 2017). In the Internet Plus era, CSR is taking on new fronts and challenges. Gamified charity is corporate philanthropy and essentially belongs to cause-related marketing, operated under the "enterprise building platform and multisocial participation" mechanism. It offers a good way for the public to join charity (public benefits), which is taking new heights in terms of quality, interactivity, and excitement. Typical examples of gamified charities are discussed below.

Recently, Alipay's "Ant Forest" and Toutiao's "Fulfilling *Dream Elf*" can be captioned as philanthropic gamified versions with Chinese characteristics. In such ICT-based environments, users become involved in charity by engaging in given tasks or activities in a virtual (gamified) setting, allowing users to perform actions and transform them into real-world scenarios. Specifically, in *Ant Forest*, users accumulate green energy from daily pro-environmental activities. When users achieve a certain amount, they can act toward a realized value (users plant virtual trees on the virtual platform, and then real trees will be planted in a specific geographic location by the platform provider) (Yang et al., 2018; Zhang et al., 2020). The game rule is that users are supposed to collect green energy manually by touching their mobile phone screens before it expires. Attributes are also offered to share the obtained green energy with friends or take it from others while having certain game rules and allowing for the planting of trees with other friends' help (Wang and Yao, 2020). In addition, the leaderboards show the achievements of users with other competitors in a gamified environment. Moreover, users can receive electronic certificates when they grow trees successfully. Users can also see real trees that are user-grown by using the map function at any time.

Fulfilling Dream Elf has a similar game design element. If users have completed specific tasks or activities relevant to the usage of Toutiao, such as daily logins, reading papers, watching videos, searching, or using other functions, the interface of *Fulfilling Dream Elf* generates a certain number of energy points. Users must collect green energy manually before it expires. When users accumulate a certain number of energy points, they can donate books in a virtual setting and transform them into real ones for underprivileged children with the help of the application provider (Toutiao).

THEORETICAL FRAMEWORK AND HYPOTHESIS DEVELOPMENT

Gamified charity is an emerging corporate philanthropy approach, which is based on digital space and presented in the form of gamification. The operation mechanism is "Enterprise building platform and multisocial participation." To some extent, gamified charities inherit the fundamental characteristics of social network games (SNGs). Various studies on SNGs across differential theoretical frameworks exist. For instance, Hsu and Lu (2004) extended the Technology Acceptance Model to predict the acceptance of online games by users. Shin and Shin (2011) integrated cognitive and affective attitudes as the acceptance model of SNGs. Wei and Lu (2014) examined the intention of mobile social games on the basis of network externalities and theory of uses and gratifications (U&G). Huang et al. (2016) also integrated U&G and customer engagement as a theoretical stance. Baabdullah (2018, 2020) revised the modified Unified Theory of Acceptance and Use of Technology as a conceptual model to explain the intention to use mobile SNGs. Moreover, Liu et al. (2016) adopted cognitive dissonance theory to explain how the dynamics of abundance of choices alter consumers' perceptions of their current e-service choices. However, all of the studies discussed above have interpreted the intention of using SNGs from different perspectives. According to Bagozzi (2007), the above-discussed theoretical stance still restricts the intervention of external variables and computation complexities. The challenging view of Bagozzi (2007) escalates its complexity, especially in the case of Information System (IS) research where virtual environment-based constructs are considered for SNGs.

In terms of human–computer interaction, interactivity (as an attribute) can be captioned as the dominating factor to define the attractiveness, involvement, and human cognition of users (Liu, 2017). Therefore, interactivity can also be counted as a construct to map the well-being of users during the experience of human–machine interaction (de Bellis and Johar, 2020). In the gamified environment, increased enjoyment (as a core of the flow experience) defines the cognitive performance of users and improves their prosocial trait and behavior. Apart from the role of immersive experience in terms of value and experience, its influence on human cognition as a research area is still evolving in the literature (i.e., virtuous cycle) (Shin, 2018), thereby heightening users' experience, related consumption, and related decision processes (Hibbeln et al., 2017). Shin (2017) concluded that the usability and learnability through immersion are directly influenced by their affordance. Similarly, Zhu et al. (2015) and Shin et al. (2016) argued that the perceived value of users and the interactive features of immersive media define their acceptance and usability. D'Errico et al. (2019) emphasized that the virtual world and its high pace of evolution have transformed the lives and perspectives of humans in recent years. For example, the affective relevance of the digital revolution has drastically revolutionized human-computer interaction as it offers affective

and empathic interfaces and digital solutions for humans. Within the pool of literature on immersive media interactivity, D'Errico et al. (2020) stated that a high degree of interactivity minimizes the psychological distance between the user and machine. Such a degree also helps to further amplify emotional arousal during the interactive experience. In addition, cognition is often referred to as perspective taking, which allows an observer to extend the affective state, emotional regulation, and prosocial behavior in the virtual space (Papapicco, 2020).

To achieve the primary objectives of this study, the persuasive behavioral and psychological view is adopted. Specifically, the Stimulus Organism Response (SOR) framework by Mehrabian and Russell (1974) is employed, which has been intensively used to explain and investigate the interaction between the environment (external factors), cognition (internal characteristics), and behavioral responses. Jacoby (2002) introduced S–O–R in the field of consumer behavior. The S–O–R framework consists of three basic components, namely, stimulus, organism, and response. Specifically, *stimulus* refers to environmental cues or external elements that influence an individual whose emotional, cognitive, or mental state of *organism* is aroused; it further takes special responsibility as a part of an internal cognitive or a psychological activity (Lin and Lo, 2016; Kamboj et al., 2018). The S–O–R framework has been widely used to research and predict behaviors while mapping entrepreneurial behavior (Liu, 2018), virtual reality tourism behavior (Kim et al., 2018), branding co-creation behavior (Kamboj et al., 2018), and knowledge-sharing behavior (Jia and Xiong, 2020). Moreover, Tang et al. (2019) explored green behavior from the SOR perspective. Participating charity behavior can also be explained by the framework (Qian et al., 2019).

Stimulus

In the context of the S-O-R framework, exogenous factors stimulate internal cognitive or psychological activities (organism), leading to a certain response. This research proposes three prominent gamified charity characteristics (*ACC*, *TEL*, and *TRA*) as stimuli. These features offer constructs driven from the application environment to predict gamified charity participatory intention (*GCPI*).

Accessibility refers to the degree of ease or convenience with which individuals can access information (Park et al., 2009; Al-debei, 2014). According to existing literature, *ACC* can be captioned as a dominant determinant that impacts the adoption of information systems or application programs, such as the satisfaction of e-banking (Liébana-Cabanillas et al., 2013), the acceptance of the digital library system (Park et al., 2009), and behavioral intention to use e-learning (Revythi and Tselios, 2019). On the contrary, the lack of *ACC* significantly influences the adoption of massive open online courses (Ma and Lee, 2018). Furthermore, Qian et al. (2019) indicated that *ACC* is a critical factor for participating in charitable activities (microcharity). Hence, the study proposes the following hypotheses:

> Hypothesis 1 (H1) (a and b): *ACC* influences user *WG* and *AR* in a gamified environment.

Telepresence can be defined as "the *experience of presence in an environment using a communication medium*," which was coined by Steuer (1992). Presence means the natural perception of an environment, while *TEL* emphasizes the sense of "being there" in the phenomenal environment created by a medium (Kim and Biocca, 1997). Fiore et al. (2005) indicated that *TEL* can be understood as an immersive response in which users experience the artificial environment provided with the necessary cognitive or sensory input. Mollen and Wilson (2010) redefined *TEL* as "*a psychological state of being there in a computer-mediated environment, enhanced by focused attention.*" Liu et al. (2020) mentioned that individuals perceive authentic experience from the mediated environment.

Although not a kind of physical or face-to-face presence, *TEL* has been accepted as equivalent to that of direct, embodied experience (Hutchins, 2011). Éric Pelet et al. (2017) also presented that *TEL* is a subjective feeling of immersion in a mediated environment, as a specific immersion case. *TEL* can essentially eliminate temporal or spatial restrictions in a large way. Thus, it is adopted in various business scenarios or product strategies, such as in an online retailer setting (Fiore et al., 2005; Suh and Chang, 2006), e-commerce setting (Lim and Ayyagari, 2018), in the hospitality industry (Ongsakul et al., 2019), social media use (Éric Pelet et al., 2017), and video games (Liu et al., 2020). Moreover, Algharabat et al. (2018) discovered that *TEL* impacts the brands of non-profit organizations, which promotes electronic word of mouth and willingness to donate. Thus, the study hypothesizes that *TEL* in gamified charities affects *WG* and *AR of the participants.*

> Hypothesis 2 (H2) (a and b): *TEL* influences user *WG* and *AR* in a gamified environment.

Meanwhile, *TRA* addresses the degree of clear view about processes and system flow. From the public administration and business management perspective, *TRA* means the degree of convenience with which external stakeholders can obtain corporate information (Bushman et al., 2004; Holland et al., 2018). That is, *TRA* refers to the openness flow of information (Bernstein, 2012). Schnackenberg and Tomlinson (2016) proposed that *TRA* entails disclosure, clarity, and accuracy. It can also be maximized by improving visibility and access to open information, great truthfulness, and information accuracy, and reducing information concealment (Yang, 2018).

The literature argues that *TRA* can be persuasive in nature to readers, especially in a philanthropic context where the general evaluation criterion is the degree of transparency. The credibility of a non-profit organization can be promoted by information transparency (Lovejoy and Saxton, 2012). Qian et al. (2019) proved that information transparency is a strategic tool for non-profit organizations to develop public relations. In addition, Cabedo et al. (2018) expressed that great transparency can help the rest of the stakeholders (not only donors) to appreciate the extent to which corporate initiatives contribute to achieving the mission of an organization. Therefore, the study proposes the following hypotheses:

Hypothesis 3 (H3) (a and b): *TRA* influences user *WG* and *AR* in a gamified environment.

Organism and Response

According to Russell (2009), the psychological state can be described as an effect, an emotion, a mood, and a feeling. He also coined the core effect as a subjective feeling and comprises two underlying dimensions of pleasure–displeasure and activation–deactivation, which can be fused in an integral whole. The structural description of the core effect explicitly matches two dichotomies of internal states in the S–O–R framework: pleasure and arousal–non-arousal (Song et al., 2021). *GCPI* can be divided into motivation-driven (*WG*) and consequence-driven (*AR*) factors as cognitive aspects.

Motivation for prosocial and pro-environmental behaviors can be aroused by the internal affections of individuals, such as seeking pleasure and pursuing happiness (Hartmann et al., 2017). Moral satisfaction can kindly drive motivation, namely, *WG* (Isen, 1970). *WG*, which originated in economics, has been adopted in various fields (e.g., green consumption) (Giebelhausen et al., 2016). It means that subjects can obtain emotional rewards by prosocial and pro-environmental behaviors (Andreoni, 1990, 1995; Taufik et al., 2015; Giebelhausen et al., 2016; Welsch et al., 2021). Jia and van Der Linden (2020) indicated that *WG* in society can drive green behavior. Moreover, Hartmann et al. (2017) found that *WG* significantly affects prosocial behavior more than altruistic value because *WG* is a psychological reward and is readily accepted by adults.

Hypothesis 4 (H4): *WG* influences user *GCPI*.

The degree of prosocial behavior of an individual is usually affected by the degree of their morality. *AR* can be interpreted as individuals who spontaneously lean toward taking responsibility for the consequences of their behaviors (Schwartz, 1977). *AR* is an integral part of NAM proposed by Schwartz (1977), which underlines the feeling of responsibility for adverse consequences caused by following non-prosocial behavior (Bamberg and Schmidt, 2003; De Groot and Steg, 2009; Klöckner, 2013; Han et al., 2020).

Many researchers have confirmed the significant relationship between *AR* and prosocial or pro-environmental behavior; for example, automobile usage behavior (Bamberg and Schmidt, 2003), transport mode change (Nordfjærn et al., 2019), energy consumption practice (Xu et al., 2020), and related pro-environmental behavior (He and Zhan, 2018; Verma et al., 2019; Han et al., 2020).

Hypothesis 5 (H5): *AR* influences the user's *GCPI*.

Mediating Role of the Organism

From the perspective of S-O-R, the organism provides a justification for bringing together stimulus and response. In this study, the affective, emotional, cognitive, or mental state of an organism is considered. Mainly, *WG* refers to emotional rewards by prosocial and pro-environmental behaviors (Andreoni, 1990), whereas *AR* reflects the feeling of responsibility for adverse consequences, potentially the outcome of non-prosocial behavior (Schwartz, 1977). The literature references *WG* as endogenous in the cases of *ACC* (Abbott et al., 2013), *TEL* (Hartmann and Apaolaza-Ibáñez, 2012), and *TRA* (Lin et al., 2017). Moreover, *WG* has been investigated while underlining its connection with the pro-environmental intention (Giebelhausen et al., 2016). Similarly, *AR* has also been observed, while emphasizing its connection with the proposed *stimulus* (Nordfjærn et al., 2019). As the mediating effect of attitude, *AR*, and personal moral norm is uncovered (Han et al., 2020), the study proposes that *WG* and *AR* have a mediating role in examining the relationship between the presented stimulus and *GCPI*.

Hypothesis 6 (H6): *WG* mediates the relationship between the proposed set of stimuli (*ACC*, *TEL*, and *TRA*) and the GCPI of users.

Hypothesis 7 (H7): *AR* mediates the relationship between the proposed set of stimuli (*ACC*, *TEL*, and *TRA*) and the GCPI of the users.

Moderating Role of NS

From the perspective of institutional theory, the behaviors of individuals and organizations can be influenced by a surrounded environment (Sambharya and Musteen, 2014). The normative pillar is a crucial attribute of institutional theory, apart from regulatory and cognitive aspects (Scott, 1995). *NS* can be labeled as "organizational and individual behavior guiding" (Bruton et al., 2010). That is, normative systems are constituted by societal values (what is admired or attractive) and societal norms (what behavior is socially acceptable or how things are done), which generate fundamental rules that individuals and organizations should conform and recognize (Valdez and Richardson, 2013). *NS* reflects specific social values in a given society. In the context of collectivist cultures, individuals tend to respond positively to their community and be easily affected by others, especially in the case of trendsetters, opinion leaders, and prominent members of society (Furnham et al., 2012; Shi et al., 2017).

The literature argues that *NS* has been adopted in several research fields, such as technology adoption (Teo et al., 2003; Liang et al., 2007), green innovation (Aguilera-Caracuel and Ortiz-de-Mandojana, 2013; Berrone et al., 2013), and entrepreneurial activity (Valdez and Richardson, 2013). Moreover, Urban and Kujinga (2017) suggested that *NS* significantly influences social entrepreneurship intentions, which are further examined in green initiatives and pro-environmental concerns by Sreen and Gleim (2020).

Hypothesis 8 (H8) (a and b): *NS* moderates the relationship between the proposed organism (*WG* and *AR*) and the GCPI of the users.

METHODOLOGY

To achieve the goal of the proposed model, we collected data with the help of an online data collection portal. We approached potential respondents through social networking sites (WeChat). The research analyzed the intention toward public willingness to participate in philanthropy (public benefit) electronically. The

study proposed the role of attributes of mobile applications (*ACC*, *TEL*, and *TRA*), which can be classified as attributes of human–machine interaction in the gamified environment to define its impact on human cognition (*AR* and *WG*) (H1–3) and further its effect on the intention to use such a digital environment for philanthropic charity (H4 and H5). In addition, the research hypothesized the role of *NS* while mapping the intention of users to use such a virtual environment. This section discusses the instruments adopted, data collection, and analyses.

Instrument

To address the instruments and the internal and external validity of the study, the instruments for each proposed construct were adopted from the existing literature. Specifically, *ACC*, as a construct, was adopted from the study of Park et al. (2011). Three item constructs for *TEL* were taken from Zhao et al. (2020). Moreover, the three-item constructs for *TRA* were adopted from the research work of Asmi et al. (2019). To examine the organism as a part of S–O–R, *WG* and *AR* were examined on the basis of the three element constructs adopted from Hartmann et al. (2017) and Zhang et al. (2019), respectively. Moreover, normative support as moderator and intention as an endogenous construct were examined by adopting the three item constructs proposed by Venkatesh and Davis (2000), Wei and Lu (2014), and Urban and Kujinga (2017). Each of the construct-related items was measured using the five-point Likert scale and listed in the table below. The Chinese version of the instruments is provided in Appendix A and is backtranslated, following the suggestion of Ye et al. (2020).

Sampling and Data Collecting

After finalizing the first version of the data collection instruments, the pilot study was conducted involving 20 potential respondents recruited from a university. Members of the pilot study provided valuable suggestions. The revised instrument (questionnaire) included other control variables (e.g., age, gender, and education). The final version of the questionnaire was distributed among potential respondents through social networking sites in the third and fourth quarters of 2020. Because the research involved human participants, the instrument was reviewed and approved by the Ethical Committee of the Department of Science and Technology of Communication - University of Science and Technology of China. Each of the potential respondents communicated about the confidentiality of the information and privacy. Specifically, potential respondents approached and inquired to participate in the survey (as briefly provided in the cover letter). To increase the response rate, follow-up messages were sent to potential respondents, resulting in a total of 669 complete responses, which accounted for further analysis. The descriptive findings are listed in **Table 1**.

TABLE 1 | Descriptive profile of the data collected.

Descriptive	Detail	Frequency	Percentage
Age	Under 20	235	35.13
	20 to 30	241	36.02
	30 to 40	104	15.55
	Above 40	89	13.30
Gender	Male	418	62.48
	Female	251	37.52
Intentions to use gamified environment	Entertainment	362	54.11
	Socialization	177	26.46
	Others	130	19.43
Average time spent in a gamified environment	Less than 30 min	583	87.14
	30–45 min	86	12.86
	More than 45 min	–	–
Which charitable APP you frequently use (choose only one)	AntForest	485	68.46
	Fulfilling Dream Elf	197	29.45
	Others	14	02.09

ANALYSIS

To examine the proposed model based on the collected data, the authors employed SPSS statistics and ADANCO 2.1 to calculate the confirmatory factor analysis (CFA) and regression to calculate the impact of moderation of the proposed setting. Specifically, the partial least squares (PLS) method was adopted to test the model and hypotheses. PLS enables structural equation modeling (SEM) to test model fitness, composite model, and factor constructs in a unified manner (Henseler et al., 2016). ADANCO, as a statistical tool, provides advanced variance-based SEM support to compute loadings, significance weights, and path coefficients.

Measurement Model

The authors used ADANCO 2.1 to compute the CFA and the loadings of the item scores to determine the internal and external validity and reliability of the model and the collected data. Favorable scores of factor loadings can be seen in the current data, as loadings are recorded over the continuum of 0.906 and 0.987. To determine the fitness indices of the proposed and saturated model, the SRMR, d_{ULS}, and dG scores were calculated with the help of ADANCO. All fitness scores were recorded within the acceptable range of values, as suggested by Henseler (2017). The tabular result of the fitness indices is provided in Appendix B.

Convergence reliability was examined by analyzing the loadings of each item, the Cronbach alpha scores and the average extracted variances (AVE); its composite reliability was also considered. The results concluded that the loadings of each item were above 0.01, which is acceptable according to Hair et al. (2014). Furthermore, the lower cut in Cronbach alpha and composite reliability was examined, following the suggestion of a previous pool of literature [see, e.g., Hair et al. (2014)]. The least scores of AVEs were also calculated, which were all observed above 0.50, as recommended by Fornell and Larcker (1981). Thus, the study can be captioned as free of the risks related to the internal reliability of the constructs. The confirmatory factor analysis scores in tabular format are listed in Appendix A.

To analyze external validity, the authors adopted two approaches, as suggested by Hair et al. (2014) and Henseler et al. (2014). Specifically, heterotrait and monotrait (HTMT) were analyzed in which scores lower than 0.850 in the case of each cross-construct relation were taken as upper cut-off points of the acceptable HTMT scores, as recommended by Henseler et al. (2016). Moreover, Fornell and Larcker's (1981) approach was considered where the correlation scores of each

construct were compared with the square roots of the AVEs of the constructs. Acceptable internal reliability was observed in the case of the current investigation. Tabular results are provided in Appendix D.

Furthermore, the extracted variance was examined to avoid the multicollinearity issue, as it can challenge the external reliability of the constructs; the acceptable range of VIF scores was calculated and is shown in Appendix A. Thus, in this study there is no risk of multicollinearity. To extend the scope of instrument and construct validity, the common method biases were calculated following Harman's single-factor analysis, as advised by Podsakoff et al. (2003) who suggested examining the maximum variance extracted by a single construct. In the present study, the maximum variance observed by a single factor was not more than 23.52%, which can help exclude the risk of common method biases.

Proposed Model

The results based on our proposed model are indicated in **Figure 1**. The variance extracted in the cases of *WG*, *AR*, and *GCPI* was observed as 61.6, 55.2, and 48.2%, respectively. The findings revealed that *ACC* has a significant positive impact on *WG* [H1(a): $\beta = 0.198$; $\rho \leq 0.05$] and *AR* [H1(b): $\beta = 0.169$; $\rho \leq 0.05$]. A similar trend can be observed in the studies by Qian et al. (2019) and Trond et al. (2019). While examining the role of *TEL*, the findings suggest that it can be captioned as the strongest stimulus among the proposed constructs. In particular, the explanatory power, while defining *WG* and *AR*, was observed as [H2(a): $\beta = 0.381$; $\rho \leq 0.001$] and [H2(b): $\beta = 0.450$; $\rho \leq 0.001$], respectively. This result supports the findings recorded by Lewis et al. (2016) and Holdack et al. (2020).

In contrast to other proposed stimuli, *TRA* was observed to be a moderate influencer while highlighting its impact on *WG* [H3(a): $\beta = 0.331$; $\rho \leq 0.001$] and *AR* [H3(b): $\beta = 0.231$; $\rho \leq 0.01$], as illustrated in **Figure 1**. The findings are in line with the results obtained by Cabedo et al. (2018). The impacts of *WG* and *AR* as exogenous factors to underline *GCPI* were recorded as significant. Specifically, the impact of *WG* (H4: $\beta = 0.497$; $\rho \leq 0.001$) was observed to be more prominent than that of *AR* (H5: $\beta = 0.255$; $\rho \leq 0.01$). However, none of the control variables (age, gender, and education) were recorded as significant. The results are listed in **Table 2**.

To examine the mediation impact of *WG* and *AR* while underlining the relationship between stimuli (*ACC*, *TEL*, and *TRA*) and their relationship to *GCPI*, the method suggested by Preacher and Hayes (2008) was adopted. Specifically, the PROCESS macro was used to perform the mediation analysis. Bootstrapping was also performed with a sample size of 10,000 to compute the asymmetric confidence intervals. The tabular result of the bootstrapping can be seen in **Table 3**. Specifically, complete mediation was observed in the case of *TRA*. Furthermore, partial mediation was recorded in the cases of *ACC* and *TEL*.

To analyze the moderation effect of *NS* in the cases of *WG* and *AR* while defining *GCPI*, hierarchal regression analysis was performed. The findings concluded that *NS* strengthens the relationship between *WG* and *GCPI* [H8(a): $\beta = 0.119$; $\rho \leq 0.001$]. However, a non-significant impact was recorded while underlining the relationship between *AR* and *GCPI* [H8(b): $\beta = 0.074$; $\rho \geq 0.05$]. The graphical interaction of the moderating effect is shown in **Figure 2**. In addition, the tabulation details are listed in Appendix C.

DISCUSSION AND IMPLICATION

Our descriptive findings conclude that the gamified setting is attractive among young people and is mostly perceived by men as trendy. Most gamified charity participants use such an application environment for entertainment and excitement

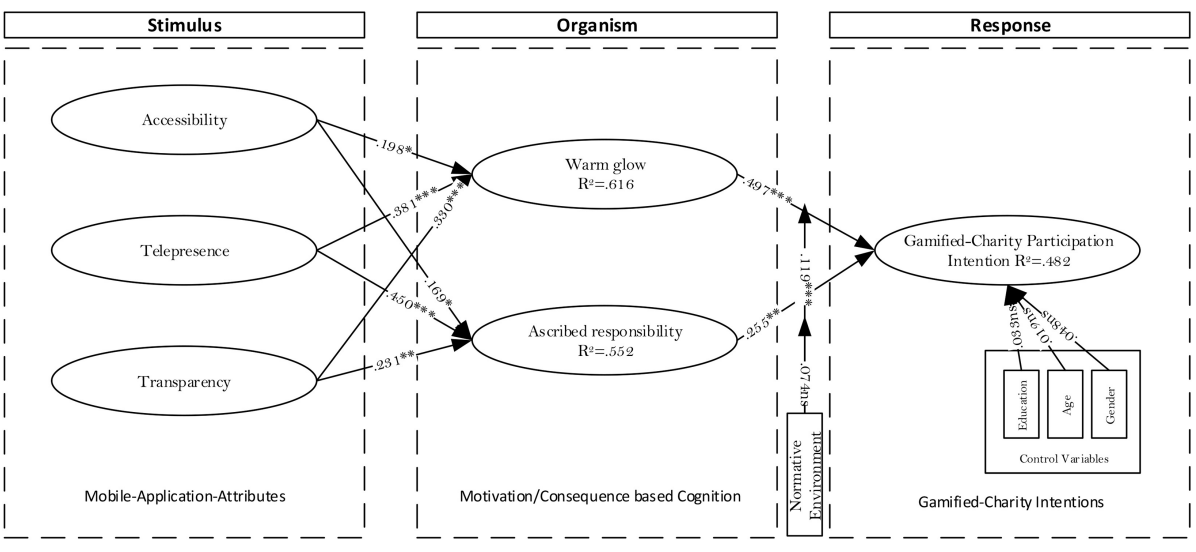


FIGURE 1 | Path analysis of the proposed model.

TABLE 2 | Path analysis.

	Effect	Original coefficient	Standard bootstrap results				Cohen's f^2	Effect size
			Mean value	Standard error	*t*-value	*p*-value (2-sided)		
H1(a)	ACC – > WG	0.198	0.196	0.107	1.847	0.044	0.065	Medium
H1(b)	ACC – > AR	0.169	0.170	0.082	2.045	0.041	0.041	Weak-medium
H2(a)	TEP – > WG	0.381	0.378	0.104	3.634	0.000	0.180	Medium
H2(b)	TEP – > AR	0.450	0.446	0.099	4.535	0.000	0.215	Medium-Strong
H3(a)	TRA – > WG	0.330	0.332	0.115	2.853	0.004	0.147	Medium
H3(b)	TRA – > AR	0.231	0.229	0.083	2.754	0.006	0.062	Medium
H4	WG – > GCPI	0.497	0.499	0.103	4.824	0.000	0.265	Medium-strong
H5	AR – > GCPI	0.255	0.252	0.092	2.750	0.006	0.069	Medium

TABLE 3 | Mediation analysis (bootstrapping results).

Hypo	IV	M	DV	Effect of IV on M	Effect of M on DV	Direct (c')	Indirect (a*b)	Total effect (c)	95% (CI)	Mediation
H6(a)	ACC	WG	GCPI	0.715***	0.376***	0.411***	0.269***	0.789***	(0.062, 0.466)	Partial
H6(b)	TEL	WG	GCPI	0.696***	0.358***	0.332***	0.249***	0.663***	(0.083, 0.438)	Partial
H6(c)	TRA	WG	GCPI	0.660***	0.491***	0.038ns	0.324***	0.509***	(0.158, 0.512)	Full
H7(a)	ACC	AR	GCPI	0.642***	0.172***	0.411***	0.110***	0.789***	(0.039, 0.274)	Partial
H7(b)	TEL	AR	GCPI	0.658***	0.125***	0.332***	0.083***	0.663***	(0.036, 0.201)	Partial
H7(c)	TRA	AR	GCPI	0.574***	0.257***	0.038ns	0.147***	0.509***	(0.035, 0.304)	Full

*** = *Significance level of 0.001.*

seeking. The research comprises several theoretical and practical implications, which are important to discuss in detail. The empirical results highlight the critical role of immersive presence and transparency of gamified charity and execution of philanthropic claims offered by the application provider while examining user acceptance and adoption, as suggested by previous literature (Davis et al., 1992; Viswanath, 2000). The literature argues that it has significant potential to predict the willingness of users to transform efforts from the potential value to the realized value (Lewis et al., 2016). However, none of the previous literature underlined *TEL* as an explanatory power simultaneously on motivation-driven (*WG*) and consequence-driven (*AR*) factors. Therefore, the study incorporated the NAM-based S-O-R framework, where stimuli based on attributes of the mobile application were considered. This aspect can be claimed as a unique contribution to IS research.

Furthermore, *TRA* has been considered an exogenous factor to map synchronization / harmonization between critical stakeholders in corporate settings (Holland et al., 2018). *TRA* also affects users in technology adoption and rejection (Mensah, 2018). However, it is never used as an exogenous factor to underline the intention of individuals toward gamified charity. Therefore, our research determined whether *TRA* would influence public intention in gamified charity, which potentially maximizes the effectiveness of corporate philanthropy in the Internet Plus era. In addition, *TRA* is noted as a fully mediated construct by *WG* and *AR*, which enriches the role of *TRA* and highlights the need to examine it further. That is, the proposed model emphasizes the strategic role of *TRA* in the case of a philanthropic application environment as it has dynamic behavior in the motivation-driven (*WG*) and consequence-driven (*AR*) factors.

In addition, *NS* strengthens the relationship between *WG* and *GCPI*. The result confirms that normative pressure can be classified as a strong determinant for restricting or guiding individuals' behaviors in the context of gamified charity, which can be claimed following previous research (Urban and Kujinga, 2017; Sreen and Gleim, 2020). By contrast, a non-significant relationship is observed between *AR* and *GCPI*, which can further lead to the gap in public understanding and framing issues. Therefore, smart, holistic, and realistic communication strategies are required. In a quantified manner, communication strategies can adopt immersive media (media-rich environment) to communicate and use real-life scenarios, which can help form effective framing strategies where society can quantify their carbon footprints more realistically than before.

Furthermore, gamified charity is corporate philanthropy and essentially belongs to cause-related marketing, operated under the "enterprise building platform and multi-social participation" mechanism, which is trendy, especially among the young generation in China. From a public perspective, gamified charity is a modern way to join philanthropy. It can also be regarded as the adoption of a new Internet application program and the participation of a recent style-based prosocial activity. Our research constructed an integrated model on the basis of the SOR framework to map public intention in gamified charities, which can be considered as a novel contribution in terms of the application of the theoretical framework.

Environment communication research argues that several initiatives struggle to communicate due to the deficiency in the message (Bertolotti and Catellani, 2014), medium (Ettinger et al., 2021; White, 2021) or communication style (Villar, 2021). In the context of risk communication, for the urge to have input from citizens, an effective framing strategy

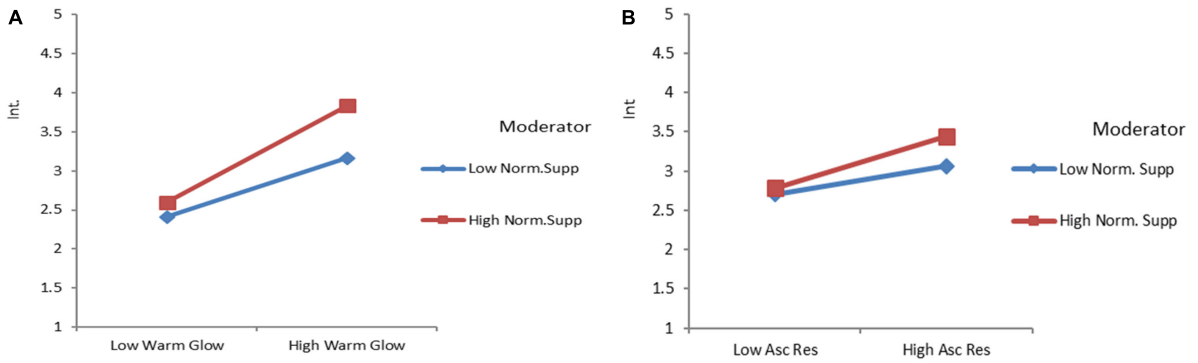


FIGURE 2 | (A) Interaction plot of Warm glow and normative support interaction plot **(B)** responsibility assignment and normative support while looking at gamified charity participation intentions.

is core to its success (Qiu et al., 2020). The current research initiative highlights the need to use a gamified environment for effective crisis communication, awareness of ecological concerns, and improvement of public ecological footprint.

Several studies have highlighted the importance of information transparency, or *TRA*. However, in our research, *TRA* was examined as an external simulative factor that led to a significant positive change in society and public behavior. Following this view, the degree of transparency of gamified charity decides public perception; thus, greater transparency can maximize public participation chances. Gamified charity has also faced severe ethical challenges, namely prosocial washing and skewed information sharing with stakeholders. Redefining the marketing communication mix is a new need for time, as a hyperactive competitive digital medium demands smart communication tools and tactics.

In the interactive / interactivity grid of users and the nature of the environment, the gamified charity can be captioned as a playful interactive setting to increase public understanding and participation and redefine perception. Our research underlines the unique participatory intention of gamified charities. It emphasizes philanthropic behavior by maximizing the sense of purpose and its relationship to potential participants. The study is individual, as it further expands the view of Marczewski's Hexad model (Tondello et al., 2016). Participants in gamified charity seem to have motivations to be philanthropists, socializers, free-spirited, achievers, and/or players. However, *ACC*, *TEL*, and *TRA* can be an integrated mobilizer to drive the change in the behavior of participants by provoking prosocial and pro-environmental values. Furthermore, as a description of *Ant Forest* and *Fulfilling Dream Elf* above, the design elements of gamified charity were observed on the basis of points/rewards, leader boards, achievements, feedback, clear goals, and narratives (Hamari and Koivisto, 2015), which can be described as the creation of social competition and the incentivizing of behavior through badge and reward systems (Hanus and Fox, 2015). Therefore, our study signifies that social attributes and reward systems can enhance the core competitiveness of *GCPI*.

Meanwhile, corporate philanthropy should adopt different geographically variated ecological concerns, such as the water crisis in a specific geographic region (water crisis can be considered), to drive philanthropic concerns and awareness. However, in another region, philanthropic issues must be addressed (air pollution or poverty problem) to drive public philanthropic intentions.

The current research environment emphasizes prosocial and pro-environmental activities operated by the "enterprise building platform and multisocial participation" mechanism, which needs legal, social and political support and attention to avoid unpleasant circumstances, such as public distrust in these initiatives. To strengthen the value impact of such initiatives, acknowledgment from macrolevel institutions is highly required.

CONCLUSION AND FUTURE RESEARCH

Gamified charity is the new paradigm of corporate philanthropy practice based on the Internet, which offers a convenient way for the public to join charity (public benefits). That is, gamified charity is formed on the basis of Chinese characteristics embedded into the DNA of Internet Plus Charity (Public Benefits) with prosocial and pro-environmental purposes. The study observes that the gamified environment plays a strategic role in the new paradigm of practice. It shows that the attributes of the philanthropic game-based application environment (including *TRA*, interactive graphical user interfaces, and *ACC*) lead to the philanthropic intentions of individuals. In addition, NAM-based factors, as part of the organism (SOR), mediate the relationship between gamified attributes and charitable intentions. However, further research can be conducted to compare the prosocial and pro-environmental attributes of the citizens in a distinctive way. The belief in efficacy and public understanding of philanthropic issues must also be addressed. Our study is unique as it underlines the significance of Chinese characteristics in participatory philanthropic intentions. These intentions in conventional (traditional) settings can be compared in future research. Corporate philanthropy, as one dimension of the CSR pyramid (Carroll, 1991), has occupied a priority position because of its high rate of visibility and return. CSR

in China is still evolving as it is mainly influenced by external actors in the supply chain network. That is, other components of the CSR pyramid should integrate with the world. Therefore, Chinese CSR practices need more attention from researchers and policymakers, especially in the case of Industry 4.0, where personal norms and CSR have a significant relationship.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The study was reviewed and approved by the Ethical Committee of the Department of Science and Technology of Communication – University of Science and Technology of China (STC-USTC).

AUTHOR CONTRIBUTIONS

RZ and ZR: concept, initialization, and funding. AS: data collection. FA and MA: data analysis. QY: write-up.

FUNDING

This work was supported by the Key Projects of Humanities and Social Science of Anhui Provincial Department of Education (Grant #: SK2018A0731) and the National Social Science Fund of China (Grant ID: 17BXW034).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2021.695468/full#supplementary-material

REFERENCES

Abbott, A., Nandeibam, S., and O'Shea, L. (2013). Recycling: social norms and warm-glow revisited. *Ecol. Econ.* 90, 10–18. doi: 10.1016/j.ecolecon.2013.02.015

Aguilera-Caracuel, J., and Ortiz-de-Mandojana, N. (2013). Green innovation and financial performance: an institutional approach. *Organ. Environ.* 26, 365–385. doi: 10.1177/1086026613507931

Al-debei, M. M. (2014). The quality and acceptance of websites: an empirical investigation in the context of higher education. *Int. J. Bus. Inf. Syst.* 15, 170–188. doi: 10.1504/ijbis.2014.059252

Alghara bat, R., Rana, N. P., Dwivedi, Y. K., Abdallah, A., and Qasem, Z. (2018). The effect of telepresence, social presence and involvement on consumer brand engagement: an empirical study of non-profit organizations. *J. Retail. Consum. Serv.* 40, 139–149. doi: 10.1016/j.jretconser.2017.09.011

Alsuwaidi, M., Eid, R., and Agag, G. (2021). Understanding the link between CSR and employee green behaviour. *J. Hosp. Tour. Manag.* 46, 50–61. doi: 10.1016/j.jhtm.2020.11.008

Andreoni, J. (1990). Impure altruism and donations to public goods: a theory of warm-glow giving. *Econ. J.* 100, 464–477. doi: 10.2307/2234133

Andreoni, J. (1995). Warm-glow versus cold-prickle: the effects of positive and negative framing on cooperation in experiments. *Q. J. Econ.* 110, 1–21. doi: 10.2307/2118508

Asmi, F., Anwar, M. A., Zhou, R., Nawaz, M. A., and Tahir, F. (2019). "Predicting functional transparency and privacy concerns as future challenge for diffusion of IOT and blockchain," in *Proceedings of the 3rd International Conferene Big Data Internet Things - BDIOT 2019*, (New York NY), doi: 10.1145/3361758.3361770

Baabdullah, A. M. (2018). Consumer adoption of mobile social network games (M-SNGs) in Saudi Arabia: the role of social influence, hedonic motivation and trust. *Technol. Soc.* 53, 91–102. doi: 10.1016/j.techsoc.2018.01.004

Baabdullah, A. M. (2020). Factors influencing adoption ofMobile social network games (M-SNGs):The role of awareness. *Inf. Syst. Front.* 22, 411–427. doi: 10.1007/s10796-018-9868-1

Bagozzi, R. P. (2007). the legacy of the technology acceptance model and a proposal for a paradigm shift. *J. Assoc. Inf. Syst.* 8, 244–254. doi: 10.17705/1jais.00122

Bamberg, S., and Schmidt, P. (2003). Incentives, morality, or Habit? Predicting Students' car use for university routes with the models of ajzen, schwartz, and triandis. *Environ. Behav.* 35, 264–285. doi: 10.1177/0013916502250134

Barnea, A., and Rubin, A. (2010). Corporate social responsibility as a conflict between shareholders. *J. Bus. Ethics* 97, 71–86. doi: 10.1007/s10551-010-0496-z

Bernstein, E. S. (2012). The transparency paradox: a role for privacy in organizational learning and operational control. *Adm. Sci. Q.* 57, 181–216. doi: 10.1177/0001839212453028

Berrone, P., Fosfuri, A., Gelabert, L., and Gomez-Mejia, L. R. (2013). Necessity as the mother of "green" inventions: institutional pressures and environmental innovations. *Strateg. Manag. J.* 34, 891–909. doi: 10.1002/smj

Bertolotti, M., and Catellani, P. (2014). Effects of message framing in policy communication on climate change. *Eur. J. Soc. Psychol.* 44, 474–486. doi: 10.1002/ejsp.2033

Bian, J., Liao, Y., Wang, Y., and Tao, F. (2021). Analysis of firm CSR strategies. *Eur. J. Oper. Res.* 290, 914–926. doi: 10.1016/j.ejor.2020.03.046

Bruton, G. D., Ahlstrom, D., and Li, H. L. (2010). Institutional theory and entrepreneurship: where are we now and where do we need to move in the future? *Entrep. Theory Pract.* 34, 421–440. doi: 10.1111/j.1540-6520.2010.00390.x

Bushman, R. M., Piotroski, J. D., and Smith, A. J. (2004). What determines corporate transparency? *J. Account. Res.* 42, 207–252. doi: 10.1111/j.1475-679x.2004.00136.x

Cabedo, J. D., Fuertes-fuertes, I., Maset-LLaudes, A., and Tirado-beltrán, J. M. (2018). Improving and measuring transparency in NGOs: a disclosure index for activities and projects. *Nonprofit Manag. Leadersh.* 28, 329–348. doi: 10.1002/nml.21298

Carroll, A. B. (1979). A Three-dimensional conceptual model of corporate performance. *Acad. Manag. Rev.* 4, 497–505. doi: 10.2307/257850

Carroll, A. B. (1991). The pyramid of corporate social responsibility: toward the moral management of organizational stakeholders. *Bus. Horiz.* 34, 39–48. doi: 10.1016/0007-6813(91)90005-g

Chen, L., and Yang, S. S. (2017). Are purchase-triggered donations advantageous with competition? *J. Oper. Res. Soc.* 68, 237–252. doi: 10.1057/s41274-016-0006-5

Davis, F. D., Bagozzi, R. P., and Warshaw, P. R. (1992). Extrinsic and intrinsic motivation to use computers in the workplace. *J. Appl. Soc. Psychol.* 22, 1111–1132.

de Bellis, E., and Johar, G. V. (2020). Autonomous shopping systems: identifying and overcoming barriers to consumer adoption. *J. Retail.* 96, 74–87. doi: 10.1016/j.jretai.2019.12.004

De Groot, J. I. M., and Steg, L. (2009). Morality and prosocial behavior: the role of awareness, responsibility, and norms in the norm activation model. *J. Soc. Psychol.* 140, 425–449. doi: 10.3200/SOCP.149.4.425-449

D'Errico, F., Leone, G., and Schmid, M. (2020). Prosocial virtual reality, empathy, and eeg measures: a pilot study aimed at monitoring emotional processes in intergroup helping behaviors. *Appl. Sci.* 10:1196. doi: 10.3390/app10041196

D'Errico, F., Paciello, M., Fida, R., and Tramontano, C. (2019). Effect of affective priming on prosocial orientation through mobile application: differences between digital immigrants and natives. *Acta Polytech. Hungarica* 16, 109–128. doi: 10.12700/APH.16.2.2019.2.7

Deterding, S., Dixon, D., Khaled, R., and Nacke, L. (2011). "From game design elements to gamefulness: defining "gamification." in MindTrek '11," in *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments (Tampere Finland: Association for Computing Machinery)*, (New York NY), 9–15.

Du, S., Bhattacharya, C. B., and Sen, S. (2010). Maximizing Business Returns To Corporate Social Responsibility (CSR): the role of CSR communication. *Int. J. Manag. Rev.* 12, 8–19. doi: 10.1111/j.1468-2370.2009.00276.x

Éric Pelet, J., Ettis, S., and Cowart, K. (2017). Information & management optimal experience of fl ow enhanced by telepresence: evidence from social media use. *Inf. Manag.* 54, 115–128. doi: 10.1016/j.im.2016.05.001

Ettinger, J., Walton, P., Painter, J., and Diblasi, T. (2021). Climate of hope or doom and gloom? Testing the climate change hope vs. fear communications debate through online videos. *Clim. Change* 164, 1–19.

Farop, O., Rupp, D., and Faroq, M. (2016). The multiple pathways through which internal and external corporate social responsibility influence organizational identification and multifoci outcomes: the moderating role of cultural and social orientations. *Acad. Manag. J.* 60, 1–69.

Fiore, A. N. N. M., Kim, J., and Lee, H. (2005). Effect of image interactivity technology on consumer responses toward the online retailer. *J. Interact. Mark.* 19, 38–53. doi: 10.1002/dir.20042

Fornell, C., and Larcker, D. F. (1981). Structural equation models with unobservable variables and measurement error: algebra and statistics. *J. Mark. Res.* 18, 382–388. doi: 10.2307/3150980

Furnham, A., Tu, B. L., and Swami, V. (2012). Cross-cultural differences in self-assessed intelligence: a comparison of british and chinese undergraduates. *Psychologia* 55, 21–27. doi: 10.2117/psysoc.2012.21

Geerts, M., and Dooms, M. (2020). An analysis of the CSR portfolio of cruise shipping lines. *Res. Transp. Bus. Manag.* 2020:100615. doi: 10.1016/j.rtbm.2020.100615

Giebelhausen, M., Chun, H. H., Cronin, J. J. Jr., and Hult, G. T. M. (2016). Adjusting the warm glow thermostat: how incentivizing participation in voluntary green programs moderates their impact on service satisfaction. *J. Mark.* 80, 56–71. doi: 10.1509/jm.14.0497

Hair, J. F., Black, W. C., Babin, B. J., and Anderson, R. E. (2014). *Multivariate Data Analysis*, 7th Edn. Upper Saddle River, N.J: Pearson Education.

Hamari, J., and Koivisto, J. (2015). Why do people use gamification services? *Int. J. Inf. Manage.* 35, 419–431. doi: 10.1016/j.ijinfomgt.2015.04.006

Han, H., Chua, B., and Ariza-montes, A. (2020). Effect of environmental corporate social responsibility on green attitude and norm activation process for sustainable consumption: airline versus restaurant. *Corp. Soc. Responsib. Environ. Manag.* 27, 1851–1864. doi: 10.1002/csr.1931

Hanus, M. D., and Fox, J. (2015). Assessing the effects of gamification in the classroom: a longitudinal study on intrinsic motivation, social comparison, satisfaction, effort, and academic performance. *Comput. Educ.* 80, 152–161. doi: 10.1016/j.compedu.2014.08.019

Hartmann, P., and Apaolaza-Ibáñez, V. (2012). Consumer attitude and purchase intention toward green energy brands: the roles of psychological benefits and environmental concern. *J. Bus. Res.* 65, 1254–1263. doi: 10.1016/j.jbusres.2011.11.001

Hartmann, P., Eisend, M., Apaolaza, V., and D'Souza, C. (2017). Warm glow vs. altruistic values: how important is intrinsic emotional reward in proenvironmental behavior? *J. Environ. Psychol.* 52, 43–55. doi: 10.1016/j.jenvp.2017.05.006

He, X., and Zhan, W. (2018). How to activate moral norm to adopt electric vehicles in China? An empirical study based on extended norm activation theory. *J. Clean. Prod.* 172, 3546–3556. doi: 10.1016/j.jclepro.2017.05.088

Henseler, J. (2017). *ADANCO 2.0.1: User Manual*. Germany: Kleve.

Henseler, J., Hubona, G., and Ray, P. A. (2016). Using PLS path modeling in new technology research: updated guidelines. *Ind. Manag. Data Syst.* 116, 2–20. doi: 10.1108/IMDS-09-2015-0382

Henseler, J., Ringle, C. M., and Sarstedt, M. (2014). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *J. Acad. Mark. Sci.* 43, 115–135. doi: 10.1007/s11747-014-0403-8

Hibbeln, M., Jenkins, J. L., Schneider, C., Valacich, J. S., and Weinmann, M. (2017). how is your user feeling? inferring emotion through human-computer interaction devices. *MIS Q. Manag. Inf. Syst.* 41, 1–21. doi: 10.25300/misq/2017/41.1.01

Holdack, E., Lurie-stoyanov, K., and Fabian, H. (2020). The role of perceived enjoyment and perceived informativeness in assessing the acceptance of AR wearables. *J. Retail. Consum. Serv.* 2020:102259. doi: 10.1016/j.jretconser.2020.102259

Holland, D., Krause, A., Provencher, J., and Seltzer, T. (2018). Transparency tested: the influence of message features on public perceptions of organizational transparency. *Public Relat. Rev.* 44, 256–264. doi: 10.1016/j.pubrev.2017.12.002

Hsu, C., and Lu, H. (2004). Why do people play on-line games? An extended TAM with social influences and flow experience. *Inf. Manag.* 41, 853–868. doi: 10.1016/j.im.2003.08.014

Huang, T., Bao, Z., and Li, Y. (2016). Why do players purchase in mobile social network games? An examination of customer engagement and of uses and gratifications theory. *Progr. Electron. Libr. Inf. Syst.* 51, 1–18.

Huotari, K., and Hamari, J. (2012). "Defining gamification-a service marketing perspective. in MindTrek'12," in *Proceeding of the 16th International Academic MindTrek Conference*, (Tampere), 17–22.

Hutchins, B. (2011). THE ACCELERATION OF MEDIA SPORT CULTURE Twitter, telepresence and online messaging. *Inform, Commun. Soc.* 14, 237–257. doi: 10.1080/1369118X.2010.508534

Isen, A. M. (1970). Success, failure, attention, and reaction to others: the warm glow of success. *J. Pers. Soc. Psychol.* 15, 294–301.

Jacoby, J. (2002). Stimulus-organism-response reconsidered: an evolutionary step in modeling (consumer) behavior. *J. Consum. Psychol.* 12, 51–57. doi: 10.1207/153276602753338081

Jia, L., and van Der Linden, S. (2020). Green but not altruistic warm-glow predicts conservation behavior. *Conserv. Sci. Pract.* 2, 1–7. doi: 10.1111/csp2.211

Jia, M., and Xiong, H. (2020). Knowledge exchange and knowledge sharing in virtual academic community: based on integrated S-O-R model and MOA theory. *Res. Libr. Sci.* 02, 43–54. doi: 10.15941/j.cnki.issn1001-0424.2020.02.007

Jia, X. (2020). Corporate social responsibility activities and firm performance: the moderating role of strategic emphasis and industry competition. *Corp. Soc. Responsib. Environ. Manag.* 27, 65–73. doi: 10.1002/csr.1774

Kamboj, S., Sarmah, B., Gupta, S., and Dwivedi, Y. (2018). Examining branding co-creation in brand communities on social media: applying the paradigm of Stimulus-Organism-Response. *Int. J. Inf. Manage.* 39, 169–185. doi: 10.1016/j.ijinfomgt.2017.12.001

Kao, E. H., Yeh, C., Wang, L., and Fung, H. (2018). The relationship between CSR and performance: evidence in China. *Pacific-Basin Financ. J.* 51, 155–170. doi: 10.1016/j.pacfin.2018.04.006

Kim, M. J., Lee, C., and Jung, T. (2018). Exploring consumer behavior in virtual reality tourism using an extended stimulus-organism-response model. *J. Travel Res.* 59, 69–89. doi: 10.1177/0047287518818915

Kim, T., and Biocca, F. (1997). Telepresence via television: two dimensions of telepresence may have different connections to memory and persuasion. *J. Comput. Commun.* 3, 1–24.

Klöckner, C. A. (2013). A comprehensive model of the psychology of environmental behaviour—A meta-analysis. *Glob. Environ. Chang.* 23, 1028–1038. doi: 10.1016/j.gloenvcha.2013.05.014

Koivisto, J., and Hamari, J. (2014). Demographic differences in perceived benefits from gamification. *Comput. Human Behav.* 35, 179–188. doi: 10.1016/j.chb.2014.03.007

Koivisto, J., and Hamari, J. (2019). The rise of motivational information systems: a review of gamification research. *Int. J. Inf. Manage.* 45, 191–210. doi: 10.1016/j.ijinfomgt.2018.10.013

Lewis, B. A., Williams, D. M., Frayeh, A. L., and Marcus, B. H. (2016). Self-efficacy versus perceived enjoyment as predictors of physical activity behavior. *Psychol Heal.* 31, 456–469. doi: 10.1080/08870446.2015.1111372.Self-Efficacy

Liang, H., Saraf, N., Hu, Q., and Xue, Y. (2007). Assimilation of enterprise systems: the effect of institutional pressures and the mediating role of top management. *MIS Q.* 31, 59–87.

Liébana-Cabanillas, F., Muñoz-Leiva, F., and Rejón-Guardia, F. (2013). The determinants of satisfaction with e-banking. *Ind. Manag. Data Syst.* 113, 750–767. doi: 10.1108/02635571311324188

Lim, J., and Ayyagari, R. (2018). Computers in human behavior investigating the determinants of telepresence in the e-commerce setting. *Comput. Human Behav.* 85, 360–371. doi: 10.1016/j.chb.2018.04.024

Lin, J., Lobo, A., and Leckie, C. (2017). The role of bene fi ts and transparency in shaping consumers' green perceived value,self-brand connection and brand loyalty. *J. Retail. Consum. Serv.* 35, 133–141. doi: 10.1016/j.jretconser.2016.12.011

Lin, S., and Lo, L. Y. (2016). Evoking online consumer impulse buying through virtual layout schemes. *Behav. Inf. Technol.* 35, 38–56.

Liu, C. (2017). A model for exploring players flow experience in online games. *Inf. Technol. People* 30, 1–60.

Liu, J. (2018). *Research on the Impact of Entrepreneurial Supports on Entrepreneurial Intentions of Postgrduate In China*. Hefei: University of Science and Technology of China.

Liu, X., Lo, V., and Wei, R. (2020). Violent videogames, telepresence, presumed influence, and support for taking restrictive and protective actions. *SAGE Open* 10, 1–11. doi: 10.1177/2158244020919524

Liu, Y., Li, H., Xu, X., Kostakos, V., and Heikkilä, J. (2016). Modeling consumer switching behavior in social network games by exploring consumer cognitive dissonance and change experience. *Ind. Manag. Data Syst.* 116, 1–21.

Lovejoy, K., and Saxton, G. D. (2012). Information, community, and action: how nonprofit organizations use social media. *J. Comput. Commun.* 17, 337–353. doi: 10.1111/j.1083-6101.2012.01576.x

Ma, L., and Lee, C. S. (2018). Investigating the adoption of MOOCs: a technology-user-environment perspective. *J. Comput. Assist. Learn.* 35, 1–10. doi: 10.1111/jcal.12314

Matten, D., and Moon, J. (2008). "Implicit" and "Explicit" CSR: a conceptual framework for a comparative understanding of corporate social responsibility. *Acad. Manag. Rev.* 33, 404–424.

Mehrabian, A., and Russell, J. A. (1974). *An Approach to Environmental Psychology*. Cambridge MA: MIT Press, 62–65.

Mensah, I. K. (2018). E-Government services adoption: the important elements of trust and transparency. *Int. J. Electron. Gov. Res.* 14, 12–31. doi: 10.4018/IJEGR.2018070102

Mollen, A., and Wilson, H. (2010). Engagement, telepresence and interactivity in online consumer experience: reconciling scholastic and managerial perspectives. *J. Bus. Res.* 63, 919–925. doi: 10.1016/j.jbusres.2009.05.014

Nan, X., and Heo, K. (2013). Consumer responses to Corporate Social Responsibility(CSR) Initiatives: examining the role of brand-cause fit in cause-related marketing. *J. Advert.* 36, 63–74. doi: 10.2753/JOA0091-3367360204

Nordfjærn, T., Lind, H. B., Şimşekoğlu, Ö, Jørgensen, S. H., Lund, I. O., and Rundmo, T. (2019). The role of social cognition in perceived thresholds for transport mode change. *Transp. Policy* 83, 88–96. doi: 10.1016/j.tranpol.2019.09.006

Ongsakul, V., Ali, F., Wu, C., Duan, Y., Cobanoglu, C., and Ryu, K. (2019). Hotel website quality, performance, telepresence and behavioral intentions. *Tour. Rev.* 76, 681–700. doi: 10.1108/TR-02-2019-0039

Papapicco, C. (2020). Google mini: italian example of artificial prosociality. *Online J. Commun. Media Technol.* 10, 1–11.

Park, N., Roman, R., Lee, S., and Chuang, J. E. (2009). User acceptance of a digital library system in developing countries: an application of the Technology Acceptance Model. *Int. J. Inf. Manage.* 29, 196–209. doi: 10.1016/j.ijinfomgt.2008.07.001

Park, S. Y., Nam, M., and Cha, S. (2011). University students' behavioral intention to use mobile learning: evaluating the technology acceptance model. *Br. J. Educ. Technol.* 43, 592–605. doi: 10.1111/j.1467-8535.2011.01229.x

Podsakoff, P. M., MacKenzie, S. B., Lee, J. Y., and Podsakoff, N. P. (2003). Common method biases in behavioral research: a critical review of the literature and recommended remedies. *J. Appl. Psychol.* 88:879.

Preacher, K. J., and Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behav. Res. Methods* 40, 879–891. doi: 10.3758/BRM.40.3.879

Qian, L., Du, L., and Hou, J. (2019). Research on the micro charity characteristic and its influences on individual charity participation behavior. *J. Manag. Sci.* 32, 120–134.

Qiu, H., Weng, S., and Wu, M. S. (2020). The mediation of news framing between public trust and nuclear risk reactions in post-Fukushima China: a case study. *J. Risk Res.* 24, 167–182. doi: 10.1080/13669877.2020.1749116

Revythi, A., and Tselios, N. (2019). Extension of technology acceptance model by using system usability scale to assess behavioral intention to use e-learning. *Educ. Inf. Technol.* 24, 2341–2355.

Robson, K., Plangger, K., Kietzmann, J. H., and Mccarthy, I. (2015). Is it all a game? Understanding the principles of gamificatio. *Bus. Horiz.* 58, 411–420. doi: 10.1016/j.bushor.2015.03.006

Russell, J. A. (2009). Emotion, core affect, and psychological construction. *Cogn. Emot.* 23, 1259–1283. doi: 10.1080/02699930902809375

Sambharya, R., and Musteen, M. (2014). Institutional environment and entrepreneurship: an empirical study across countries. *J. Int. Entrep.* 12, 314–330. doi: 10.1007/s10843-014-0137-1

Sardi, L., Idri, A., and Fernández-alemán, J. L. (2017). A systematic review of gamification in e-Health. *J. Biomed. Inform.* 71, 31–48. doi: 10.1016/j.jbi.2017.05.011

Schnackenberg, A. K., and Tomlinson, E. C. (2016). Organizational transparency: a new perspective on managing trust in organization-stakeholder relationships. *J. Manage.* 42, 1784–1810. doi: 10.1177/0149206314525202

Schwartz, S. H. (1977). Normative influences on altruism. *Adv. Exp. Soc. Psychol.* 10, 221–279. doi: 10.1016/S0065-2601(08)60358-5

Scott, W. R. (1995). *Institutions and Organizations. Ideas,Interests and Identities*. Thousand Oaks, CA: SAGE Publications Ltd.

Shi, H., Wang, S., and Zhao, D. (2017). Exploring urban resident's vehicular PM2.5 reduction behavior intention: an application of the extended theory of planned behavior. *J. Clean. Prod.* 147, 603–613. doi: 10.1016/j.jclepro.2017.01.108

Shin, D. (2017). The role of affordance in the experience of virtual reality learning: tech- nological and affective affordances in virtual reality. *Telemat. Inform.* 34, 1826–1836. doi: 10.1016/j.tele.2017.05.013

Shin, D. (2018). Computers in human behavior empathy and embodied experience in virtual environment: to what extent can virtual reality stimulate empathy and embodied experience? *Comput. Human Behav.* 78, 64–73. doi: 10.1016/j.chb.2017.09.012

Shin, D., and Shin, Y. (2011). Why do people play social network games? *Comput. Human Behav.* 27, 852–861. doi: 10.1016/j.chb.2010.11.010

Shin, D., Choi, M., and Kim, J. H. (2016). Interaction, engagement, and perceived interactivity in single-handed interactionn. *Inter. Res.* 26, 1134–1157. doi: 10.1108/IntR-12-2014-0312

Song, S., Yao, X., and Wen, N. (2021). What motivates Chinese consumers to avoid information about the COVID-19 pandemic: the perspective of the stimulus-organism-response model. *Inf. Process. Manag.* 58:102407. doi: 10.1016/j.ipm.2020.102407

Sreen, N., and Gleim, M. (2020). The impact of the institutional environment on green consumption in India. *J. Consum. Mark.* 38, 47–57. doi: 10.1108/JCM-12-2019-3536

Steuer, J. (1992). Defining virtual reality: dimensions determining telepresence. *J. Commun.* 42, 73–93. doi: 10.1111/j.1460-2466.1992.tb00812.x

Su, L., and Swanson, S. R. (2019). Perceived corporate social responsibility's impact on the well-being and supportive green behaviors of hotel employees: the mediating role of the employee-corporate relationship. *Tour. Manag.* 72, 437–450. doi: 10.1016/j.tourman.2019.01.009

Suh, K., and Chang, S. (2006). User interfaces and consumer perceptions of online stores: the role of telepresence. *Behav. Inf. Technol.* 25, 99–113. doi: 10.1080/01449290500330398

Tang, Z., Warkentin, M., and Wu, L. (2019). Understanding employees' energy saving behavior from the perspective of stimulus-organism-responses. *Resour. Conserv. Recycl.* 140, 216–223. doi: 10.1016/j.resconrec.2018.09.030

Taufik, D., Bolderdijk, J. W., and Steg, L. (2015). Acting green elicits a literal warm glow. *Nat. Clim. Chang.* 5, 37–40. doi: 10.1038/NCLIMATE2449

Teo, H., Wei, K., and Benbasat, I. (2003). Predicting intention to adopt interorganizational linkages: an institutional perspective. *MIS Q.* 27, 19–49.

Tondello, G. F., Wehbe, R. R., Diamond, L., Busch, M., Marczewski, A., and Nacke, L. E. (2016). "The gamification user types hexad scale," in *Proceedings of the The ACM SIGCHI Annual Symposium on Computer-Human Interaction in Play*, (Austin, TX), 229–243. doi: 10.1145/2967934.2968082

Trond, N., Lina, H. B., Şimşekoğlu, Ö, Jørgensen, S. H., Lund, I. O., and Rundmo, T. (2019). The role of social cognition in perceived thresholds for transport mode change. *Transp. Policy* 83, 88–96. doi: 10.1016/j.tranpol.2019.09.006

Urban, B., and Kujinga, L. (2017). The institutional environment and social entrepreneurship intentions. *Int. J. Entrep. Behav. Res.* 23, 638–655. doi: 10.1108/IJEBR-07-2016-0218

Valdez, M. E., and Richardson, J. (2013). Institutional determinants of macro-level entrepreneurship. *Entrep. Theory Pract.* 37, 1149–1175. doi: 10.1111/etap.12000

Venkatesh, V., and Davis, F. D. (2000). A theoretical extension of the technology acceptance model: four longitudinal field studies. *Manage. Sci.* 46, 186–204. doi: 10.1287/mnsc.46.2.186.11926

Verma, V. K., Chandra, B., and Kumar, S. (2019). Values and ascribed responsibility to predict consumers' attitude and concern towards green hotel visit intention. *J. Bus. Res.* 96, 206–216. doi: 10.1016/j.jbusres.2018.11.021

Villar, M. E. (2021). Community engagement and co-creation of strategic health and environmental communication: collaborative storytelling and game-building. *J. Sci. Commun.* 20, 1–9.

Viswanath, V. (2000). Determinants of perceived ease of use: integrating control, intrinsic motivation, and emotion into the technology acceptance model. *Inf. Syst. Res.* 11:342.

Wang, X., and Yao, X. (2020). Fueling pro-environmental behaviors with gamification design: identifying key elements in ant forest with the kano model. *Sustain* 12, 2213. doi: 10.3390/su12062213

Wei, P., and Lu, H. (2014). Why do people play mobile social games? An examination of network externalities and of uses and gratifications. *Internet Res.* 24, 313–331. doi: 10.1108/IntR-04-2013-0082

Welsch, H., Binder, M., and Blankenberg, A. (2021). Green behavior, green self-image, and subjective well-being: separating affective and cognitive relationships. *Ecol. Econ.* 179, 106854. doi: 10.1016/j.ecolecon.2020.106854

White, T. (2021). Interactive and immersive documentary methods for community resilience & disaster reduction. *Photographies* 14, 95–107. doi: 10.1080/17540763.2020.1847177

Xiao, H., and Yang, Z. (2020). Corporate social responsibility of platform enterprises: logical origin and practical paradigm. *Bus. Manag. J.* 42, 37–53.

Xiao, H., Yang, Z., and Jiang, B. (2020). Development of corporate philanthropy in china: its development pattern and future scenario. *J. NANJING Univ.* 57, 32–50.

Xu, X., Chen, C., Li, D., Menassa, C., Mccunn, L. J., and Chen, C. (2020). Energy saving at work: exploring the role of social norms, perceived control and ascribed responsibility in different office layouts. *Front. Built Environ.* 6:16. doi: 10.3389/fbuil.2020.00016

Yang, Y. (2018). Is transparency a double-edged sword in citizen satisfaction with public service? Evidence from China's public healthcare. *J. Serv. Theory Pract.* 28, 484–506. doi: 10.1108/JSTP-06-2017-0102

Yang, Z., Kong, X., Sun, J., and Zhang, Y. (2018). Switching to green lifestyles: behavior change of ant forest users. *Int. J. Environ. Res. Public Health* 15, 1819. doi: 10.3390/ijerph15091819

Ye, Q., Anwar, M. A., Zhou, R., Asmi, F., and Ahmad, I. (2020). China's green future and Household Solid Waste: challenges and Prospects. *Waste Manag.* 105, 328–338. doi: 10.1016/j.wasman.2020.02.025

Yu, H.-C., Kuo, L., and Kao, M.-F. (2017). The relationship between CSR disclosure and competitive advantage. *Sustain. Accounting, Manag. Policy J.* 8, 547–570.

Zhang, B., Lai, K. H., Wang, B., and Wang, Z. (2019). From intention to action: how do personal attitudes, facilities accessibility, and government stimulus matter for household waste sorting? *J. Environ. Manage.* 233, 447–458. doi: 10.1016/j.jenvman.2018.12.059

Zhang, M., Tong, L., Su, J., and Cui, Z. (2015). Analyst coverage and corporate social performance: evidence from China. *Pacific-Basin Financ. J.* 32, 76–94. doi: 10.1016/j.pacfin.2015.02.005

Zhang, Y., Xiao, S., and Zhou, G. (2020). User continuance of a green behavior mobile application in China: an empirical study of Ant Forest. *J. Clean. Prod.* 242, 118497. doi: 10.1016/j.jclepro.2019.118497

Zhao, Y., Wang, A., and Sun, Y. (2020). Technological environment, virtual experience, and MOOC continuance: a stimulus–organism–response perspective. *Comput. Educ.* 144, 103721. doi: 10.1016/j.compedu.2019.103721

Zhu, Y., Heynderickx, I., and Redi, J. A. (2015). Understanding the role of social context and user factors in video quality of experience. *Comput. Hum. Behav.* 49, 412–426. doi: 10.1016/j.chb.2015.02.054

Zhou, H., Li, D., and Zhang, W. (2019). Strategic corporate social responsibility and corporate performance: can the best of both worlds be realized? *Res. Econ. Manag.* 40, 131–144.

Conflict of Interest: QY and FA were affiliated to company Anhui Xinhua Media Co., Ltd. as part-time researchers. RZ was a consultant for the company Anhui Xinhua Media Co., Ltd.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Ye, Rafique, Zhou, Asmi, Anwar and Siddiquei. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Psychology

ORIGINAL RESEARCH
published: 30 September 2021
doi: 10.3389/fpsyg.2021.615647

Shape–Trait Consistency: The Matching Effect of Consumer Power State and Shape Preference

Feng Yao[1], *Xiaotong Jin*[1], *Banggang Wu*[2]*, *Taiyang Zhao*[3] and *Tiannv Ma*[4]

[1] Department of Marketing, Business School, Jilin University, Changchun, China, [2] Department of Marketing, Business School, Sichuan University, Chengdu, China, [3] Department of Psychology, School of Philosophy and Sociology, Jilin University, Changchun, China, [4] Department of Business Administration, School of Management, Changchun University, Changchun, China

Angular and rounded shapes are two important visual elements widely used in the design of product shapes and brand logos. By introducing the power state, a psychological variable that is inherently relevant to consumers' product choices, brand preferences, and decision-making, we propose that consumers' power state influences their shape preference. Specifically, compared to low-power consumers, high-power consumers respond more positively to angular as opposed to rounded shapes, because the angular shape facilitates the expression of competence (as opposed to warmth). Through four studies, we provide consistent support for our main predictions as well as the underlying processes. Studies 1 and 2 demonstrate that consumers experiencing higher power are more likely to prefer an angular shape over a rounded shape than those experiencing lower power through different research methods, research objects, and experimental materials. On this basis, studies 3 and 4 further explore the mechanisms underlying the observed effects. These findings contribute to sensory marketing and power research and provide important implications for visual design and advertisement development.

Keywords: power state, shape preference, competence, warmth, product design

OPEN ACCESS

Edited by:
Oliver Niebuhr, University of Southern Denmark, Denmark

Reviewed by:
Siegfried Dewitte, KU Leuven, Belgium
Liying Zhou, Sichuan University, China

***Correspondence:**
Banggang Wu
wubg@scu.edu.cn

Specialty section:
This article was submitted to Personality and Social Psychology, a section of the journal Frontiers in Psychology

Received: *09 October 2020*
Accepted: *10 August 2021*
Published: *30 September 2021*

Citation:
Yao F, Jin X, Wu B, Zhao T and Ma T (2021) Shape–Trait Consistency: The Matching Effect of Consumer Power State and Shape Preference. Front. Psychol. 12:615647. doi: 10.3389/fpsyg.2021.615647

INTRODUCTION

Shape design is an important marketing element in product and brand logo design. Different shape designs communicate different product and brand connotations and also attract consumers with different psychological characteristics. Exploring consumers' shape preferences can not only help in understanding the psychological demands of consumers but also to accurately provide consumers with suitable products. The sense of power, as a basic psychological variable, has an important influence on consumers' cognition, emotion, and behavior. Therefore, when faced with different shape designs, how do consumers with different power states choose?

Power is defined as "asymmetric control over valued resources in a social relationship" (Galinsky et al., 2008, 2015). Previous research has shown the power state has an important impact on a variety of consumer behaviors (Jin and Huang, 2018). However, existing research on the power state in consumer behavior areas mainly focuses on consumer persuasion (Dubois et al., 2016; Jin and Huang, 2018), compensatory consumption (Dubois et al., 2012), and behavior related to money and price (Garbinsky et al., 2014), which rarely provide direct guidance information for product design. Therefore, this study seeks to determine the kind of product shape that can better meet the preferences of consumers in different power states. Thus, it provides guidance and suggestions for product design.

Recently, angular and rounded shapes as important visual elements in product design have received increased attention from scholars and marketers. For example, Vallen et al. (2018) demonstrated that salespersons are more likely to recommend round (vs. angular) shaped products to heavier (vs. thinner) customers. Research on product selection suggests that in the process of product selection, people will seek the congruence between self-concept and product impression (i.e., self-image congruence; Sirgy, 1982; Sirgy et al., 2004). Previous studies revealed that individuals with high or low power states have different self-concepts (i.e., how individuals regard themselves) in many aspects. For example, compared with low-power consumers, high-power consumers are more confident in themselves, more creative, and have higher executive functioning (Galinsky et al., 2015).

Therefore, consumers with different power states would choose products that match their personal characteristics. According to the agentic–communal model of power, consumers experiencing a psychological state of high power have higher demand for competence. In contrast, consumers experiencing low-power state have higher demand for warmth (Dubois et al., 2016), which parallels the characteristics of angular and rounded shapes. Specifically, the angular shape expresses competence, while the rounded shape primarily expresses warmth. Therefore, when making consumption choices related to shape, high-power consumers prefer an angular shape with competence perception, while low-power consumers prefer rounded shapes with warmth perception.

THEORETICAL FRAMEWORK

Power and Psychological Propensities

Psychological propensities of power refer to inclinations or tendencies naturally triggered by the psychological experience of power (Rucker et al., 2012). It is a behavioral tendency requiring little to no cognitive intervention. For example, power enhances people's action tendencies (Galinsky et al., 2003), making people more confident (Brinol et al., 2007). Similarly, power states affect consumers' psychological preferences for objective things (Kim and McGill, 2011).

Drawing on the seminal work of Bakan (1966), Rucker et al. (2012) proposed an agentic-communal model of power, suggesting the power state affects individuals' agentic-communal orientation and further affects individuals' thoughts and behaviors. Based on the agentic-communal model, high power tends to foster an agentic orientation where the self is viewed as more important and valued, and pursues independence and self-struggle more freely, relying less on others. Therefore, compared with low-power consumers, high-power consumers pay more attention to their own capabilities. In contrast, low power tends to foster a communal orientation where others are viewed as more important and valued (Rucker et al., 2011) and has a stronger group tendency, paying more attention to the needs of others. Due to their lack of sufficient capabilities and resources, low-power individuals often need to rely on others to achieve self-worth. Therefore, compared with high-power individuals, low-power individuals pay more attention to the connection with members of the surrounding groups and have a stronger appeal for warmth.

Meanwhile, according to the psychological tendency theory of power, individuals prefer to choose products that match their own characteristics (Galinsky et al., 2015). Thus, high-power individuals are more inclined to choose options related to competence. As one example, Rucker and Galinsky (2009) proposed that compared with low-power consumers, high-power consumers respond more positively to persuasive appeals that emphasize the quality and performance of products. On the contrary, the behavior and consumption choices of low-power individuals are more likely to be related to warmth. For example, low-power (vs. high-power) individuals are more willing to donate (Roux et al., 2012). Based on the above inferences, high-power consumers demonstrate they have a higher demand for competence, while low-power consumers have a higher demand for warmth. Dubois et al. (2016) also confirmed this view. In their research, they found high-power individuals pay more attention and are more easily persuaded by information related to competence; on the contrary, individuals with low power pay more attention and are more easily persuaded by information related to warmth. The existing research on power state and perceived competence and warmth mainly focuses on persuasion, but current research breaks through this limitation and introduces it into the field of sensory marketing to explore the relationship between power state and shape preference and its explanation mechanism. Meanwhile, through empirical test, it provides direct empirical evidence for the relationship between power state and perceived competence and warmth.

Visual Shape and Stereotype Content Model

Angular and rounded are typical representatives of shape in visual design (Vallen et al., 2018; Yang and Chen, 2019). Angular shapes (e.g., rectangles, triangles, etc.) are composed of straight lines, planes, and sharp corners, while rounded shapes (e.g., circles, ellipses, etc.) are composed of smooth curves and arcs. The difference between rounded and angular shapes is not only reflected in the composition of the elements but also the symbolic meanings they hold (Arnheim, 1974). Different shapes bring different perceptive characteristics to consumers, which play an important role in consumers' value judgment and decision-making (Zhang et al., 2006; Zhu and Argo, 2013; Jiang et al., 2016). For example, Lieven et al. (2015) show that the shape of a brand logo will affect consumers' perception of brand "personalities." Specifically, an angular shape will enhance the perception of the masculinity of brands, whereas rounded shapes will enhance the perception of the femininity of brands. Zhu and Argo (2013) demonstrated that in the negotiation process, an angular seating arrangement triggers a situational need to be unique, and individuals seated in such a setting respond more favorably to a self-oriented persuasive appeal, while a rounded seating arrangement triggers a situational need to belong and respond more favorably to a family-oriented persuasive appeal.

According to the stereotype content model (SCM; Fiske et al., 2007), individuals evaluate the target object through two

fundamental dimensions of social judgments: competence and warmth. Warmth embodies traits of being friendly, trustworthy, and kindness, whereas competence is associated with traits of being competent, efficient, and skillful (Fiske et al., 2002; Aaker et al., 2010). The evaluation of the competence and warmth for the target object will further trigger the individuals' positive or negative affect, and then drive their consumption behavior. In the current study, we introduce competence and warmth into the expression of shape "personality," recognizing the angular and rounded shape from the perspective of competence and warmth.

Prior studies on emotional expressions have revealed that people tend to use curved lines to express emotional states of gentleness, happiness, and quietness, while angled lines are used to express serious, harsh, and vigorous emotions (Lundholm, 1921; Poffenberger and Barrows, 1924). Jiang et al. (2016) also proposed that an angular brand logo leads to a sense of hardness and durability, while a rounded brand logo triggers softness and comfort perception to consumers. Therefore, based on the above research, we infer that compared to the rounded shape, the angular shape is superior to the perception of competence, while the rounded shape is superior in terms of the perception of warmth. This fully reflects the real-life symbolic meaning of the shapes. In work situations or sales services, for example, an angular face gives people a sense of seriousness and capability (competence), whereas a rounded face makes people looks gentler and friendlier (warmth).

In summary, the competence and warmth embodied by angular and rounded shapes are consistent with the appeal of high- and low-power consumers. Therefore, in the research on visual marketing related to shapes, it is proposed that the power state affects consumers' shape preference, and this influence is explained by the mediating effect of competence and warmth. Specifically, high-power individuals have higher demands for competence, so they tend to choose angular shapes that reflect competence characteristics, while low-power individuals have higher demands for warmth, so they prefer rounded shapes that reflect warmth. Accordingly, we propose the following hypotheses:

Hypothesis 1: Consumers experiencing higher power are more likely to prefer an angular shape over a rounded shape than consumers experiencing lower power.

Hypothesis 2: Perceived competence and warmth play a mediating role between power state and shape preference.

The above inferences about the mediating effect put forward the mechanisms underlying perceived competence and warmth. However, according to existing research theories, we also try to rule out two possible alternative explanations: conflict resolution types and need for uniqueness. The literature regarding angular and rounded shapes suggests angular shapes tend to generate confrontational associations, and rounded shapes tend to generate compromise associations (Zhang et al., 2006). According to the approach-inhibition theory of power (Keltner et al., 2003), having power activates consumers' approach-related tendencies, whereas lacking power activates consumers' inhibition-related tendencies. Therefore, conflict resolution types may be the mechanisms underlying the influence of power state on shape preference. Specifically, consumers with high power have a higher motivation to approach behavior, so they tend to be more confrontational, and thus prefer an angular shape. On the contrary, low-power consumers have a higher motivation to inhibit behavior and tend to prefer compromise, thus preferring a rounded shape. In addition, Zhu and Argo (2013) demonstrated that angular-shaped seating arrangements prime a need for uniqueness, while the rounded shaped prime a need to belong. Can this matching effect between shape and the need for uniqueness in seating arrangements be carried over into general research into shape preferences? Previous studies have shown that consumers with different power states will choose products that match their personal characteristics (Galinsky et al., 2015). Consumers experiencing a state of power are more likely to seek uniqueness, whereas those experiencing a state of powerlessness tend to seek conformity (Jin et al., 2011). Therefore, compared with low-power consumers, high-power consumers may have a higher need for uniqueness, and thus prefer angular (vs. rounded) shapes. The above argument preliminarily demonstrates possible alternative explanations of conflict resolution style and the need for uniqueness, but compared with perceived competence and warmth, whether they are the main factors affecting consumers' preference for angular and rounded shapes in different power states must be further empirically explored.

In the next section, we report the empirical tests of our hypotheses. Studies 1 and 2 test our basic hypothesis (H1) that consumers in high-power states indicate a more favorable attitude toward angular shapes through surveys and experiments. Studies 3 and 4 explore the mechanisms underlying the observed effects (H2) and excludes two possible alternative explanations. Study 5 further verifies the mediating effect of this research by manipulating the mediating variables.

STUDY 1: CORRELATIONAL EVIDENCE

Study 1 provides an initial test of Hypothesis 1 that consumers in high- and low-power states respond differently to angular vs. rounded shapes. We measure consumers' chronic sense of power and predicted consumers with a higher sense of power would respond more positively to an angular shape. In contrast, consumers with a lower sense of power would respond more positively to rounded shapes. Study 1 consists of two sub-studies (Study 1a and Study 1b) verifying Hypothesis 1 with participants from different sources. Study 1a was predominantly American, while Study 1b was predominantly Chinese, to ensure consistency of results across cultures.

Study 1a

Method

Sample

A total of 311 workers (155 females and 156 males; $M_{age} = 34.06$, $SD = 10.83$) from the Prolific participant pool took part in the online study. The number of participants was specified a priori based on power analysis in several pilot studies with samples from the same population. The study took ∼2 min, and each participant received a $0.3 reward for completing the study.

Procedure

We first used the scale to measure participants' general sense of power (Anderson et al., 2012). The scale consists of eight statements (e.g., *I think I have a great deal of power*) and the participants need to rate the extent to which they agreed with each statement, using a 7-point Likert scale (1 = *strongly disagree* to 7 = *strongly agree*, $\alpha = 0.92$). We then showed participants three pairs of images (see **Appendix A**). One pair of images contained simple shapes, and the other two pairs of images contained brand logos. Simple shapes were included in order to eliminate the interference of other complex factors, and brand logos were included in order to confirm the universality of our findings and provide empirical evidence for the subsequent design of actual products and brand logos. In each pair, there were two corresponding images: one with a circular shape and another with an angular shape. According to the prior test, for the three pairs of images in this study, the participants perceived the angularity of the angular graphics to be significantly higher than the rounded graphics (*ps* < 0.05). At the same time, there were no significant differences in the aesthetics of angular and rounded graphics in each pair (*ps* > 0.05). In addition, the order of presentation of each pair and the images within each pair were both randomized. In line with previous research (Zhang et al., 2006) on shape references, participants were asked to rate their preference between each pair of images (1 = I definitely prefer A to B, 7 = I definitely prefer B to A). Because the angular shape or rounded shape images were presented randomly in the position of A or B, we recoded this variable in such a way that a higher score indicated preference for an angular shape over a rounded shape.

We also collected participants' information about age and sex and analyzed them as control variables because they have been shown to covary with power and shape preference in previous studies (Lieven et al., 2015; Rucker et al., 2018). In addition, we measured participants' subjective social status (i.e., respect, admiration, and voluntary deference in the eyes of others; Magee and Galinsky, 2008). As for the measurement of subjective social status, consistent with prior research (Blader and Chen, 2012; Blader et al., 2016), participants rated the extent to which they agreed with three statements, "People respect me," "People admire me," and "People voluntary respect me," using the 7-point Likert scale (1 = *strongly disagree* to 7 = *strongly agree*; α =0.89). Although power and social status are distinct constructs (Keltner et al., 2003; Blader and Chen, 2012), they tend to be correlated and can be associated with similar outcomes (Magee and Galinsky, 2008). Thus, to show the effects of power above and beyond that of social status, social status serves as a control variable in our analysis.

Results and Discussion

In support of Hypothesis 1, the statistical results of the simple shape show the general sense of power was significantly positively correlated with preference for an angular shape over a rounded shape (Simple shape: $r = 0.12$, $p = 0.032$), which was also confirmed by the analysis of brand logos (Logo 1: $r = 0.16$, $p = 0.006$; Logo 2: $r = 0.13$, $p = 0.025$). In addition, the combined data of the three experimental image pairs showed similar results, thereby providing further evidence to verify Hypothesis 1 ($r = 0.19$, $p = 0.001$). When controlling for participants' age, sex, and social status, power remained to be significantly positively related to the preference for an angular shape over a rounded shape ($\beta = 0.19$, $p < 0.001$). Specifically, Simple shape: $\beta = 0.11$, $p = 0.046$; Logo 1: $\beta = 0.15$, $p = 0.005$; Logo 2: $\beta = 0.14$, $p = 0.012$. Thus, participants with a higher sense of general power in their lives were more likely to prefer an angular shape over a rounded shape than those who had lower power.

Study 1a chose one simple shape and two different logo graphics as experimental materials for dependent variables and verified the influence of power state on shape preference. In addition to simple shapes and logos, angular and rounded design elements can also be reflected in products. Therefore, although the procedure of Study 1b is similar to that of Study 1a, there are two main considerations in the design of Study 1b; one is the replacement of experimental materials, and the other is the cross-cultural research of participants. In Study 1b, we chose two specific products, cushion and mug, as our experimental materials (See **Appendix B**) to test whether the power state will affect the preference of product shape. Additionally, the participants of Study 1b were Chinese.

Study 1b

Method

Sample

A total of 233 adult consumers (132 females and 101 males; $M_{age} = 21.84$, $SD_{age} = 5.31$) participated in this study, and each participant received a ¥1 reward for completing the study.

Procedure

The procedure of Study 1b was consistent with that of Study 1a. Participants were asked to measure their sense of power, shape preference for the two pairs of images, and the demographic information in turn.

Results and Discussion

The experimental results also support Hypothesis 1—general sense of power was significantly positively correlated with the preference for an angular shape over a rounded shape (Cushion: $r = 0.13$, $p = 0.043$; Mug: $r = 0.17$, $p = 0.010$). In addition, as for control variables, in Study 1a, we controlled for participants' age, sex, and social status, but in Study 1b, we only controlled for participants' sex and age because most participants in this study were students. In contrast to those who have been working, students have poor sensitivity to social status. When controlling for participants' age and sex, the positive influence of power on the preference for an angular shape was still significant (Cushion: $\beta = 0.13$, $p = 0.047$; Mug: $\beta = 0.15$, $p = 0.021$).

Studies 1a and 1b measured participants' sense of power as an individual differing factor, provided initial evidence for Hypothesis 1 that consumers experiencing higher power are more likely to prefer an angular shape over a rounded shape than consumers experiencing lower power. In addition, Studies 1a and 1b used different experimental materials and research objects. The above study design and statistical results show the conclusions of this study have good external validity. To better

establish the causal relationship, in Study 2, instead of measuring the sense of power, we experimentally manipulated the high vs. low power state of participants to verify Hypothesis 1.

STUDY 2: EXPERIMENTAL EVIDENCE

Study 2 aimed to verify H1 through an experimental method. Study 2 also included two sub-studies (Study 2a and Study 2b) with participants from different countries. The participants in Study 2a were mainly American, while those in Study 2b were mainly Chinese.

Study 2a

Method

Sample and Design

A total of 149 workers (61 females and 83 males; $M_{age} = 33.11$, $SD = 9.00$) were recruited from Amazon's Mechanical Turk to participate in this study. They were randomly assigned to the conditions of a single factor 3 (power: high vs. control vs. low) between-subjects design. The study took ∼5 min, and each participant received a $0.6 reward for completing the study.

Procedure

We conducted the experiment as two unrelated studies. First, participants were asked to complete an episodic recall task (Galinsky et al., 2003; Rucker and Galinsky, 2008), frequently used in previous research to manipulate their power state. We asked participants in the high-power condition to recall a particular incident in which they had power over another individual and describe this situation in 5–7 sentences—namely what happened and how they felt. Participants in the low-power condition were asked to recall a particular incident in which someone else had power over them and describe this situation in 5–7 sentences—namely what happened and how they felt. While we asked participants in the control condition to recall a recent trip to a grocery store or convenience store, the requirements were the same as the other two conditions. Upon completion of the episodic recall task, participants were asked to rate the extent to which they feel "powerful, influential, important, powerless, subordinate, dependent" ($\alpha = 0.86$) on 7-point scales as a power manipulation check. Items including powerless, subordinate, and dependent are reverse scored for the scale calculations. Subsequently, the participants were presented with a picture of a set of graphics (Logo 2; see **Appendix C**) and asked their preference for an angular shape over a rounded shape. The detailed measurement is the same as Study 1a. After completing the survey, the participants were debriefed and thanked.

Results

Manipulation Checks

The results of ANOVA showed that the power state was successfully manipulated [$F_{(2,146)} = 33.88$, $p < 0.001$]. Specifically, participants who were assigned to the high-power condition ($M = 5.24$, $SD = 0.77$) reported more powerful than those who were assigned to the low-power condition [$M = 3.40$, $SD = 1.31$; $t_{(92)} = 8.45$, $p < 0.001$] and control condition ($M = 4.20$, $SD = 1.15$; $t_{(105)} = 5.47$, $p < 0.001$). The degree of power reported by participants in the low-power condition was significantly lower than that in the control group [$t_{(95)} = 3.17$, $p = 0.002$).

Preference

An ANOVA performed on the preference score yielded a significant influence, $F_{(2,146)} = 3.27$, $p = 0.041$. Participants in the high-power condition ($M = 5.44$, $SD = 1.72$) reported more positive attitudes toward angular shape over rounded shape than those in the low-power condition [$M = 4.45$, $SD = 2.35$; $t_{(92)} = 2.36$, $p = 0.021$] and control condition [$M = 4.65$, $SD = 2.04$; $t_{(105)} = 0.034$], while the attitude between low-power condition and control condition was not significantly different [$t_{(95)} = 0.45$, $p = 0.652$].

Study 2b

The procedure of Study 2b was similar to that of Study 2a, but with three major changes. First, instead of the online experiment, we conducted this study in the laboratory. Second, participants in this study were mainly Chinese. Third, to test the robustness of our effect, the experimental materials of the dependent variables in Study 2b were different from those of Study 2a.

Method

Sample and Design

A total of 166 undergraduate students from a public university participated in this study (89 females and 77 males; $M_{age} = 19.89$, $SD_{age} = 3.46$) and each participant received a ¥3 reward for completing the study. Participants were randomly assigned to one of the three conditions of a single factor 3 (power: high vs. control vs. low) between-subjects design.

Procedure

Participants completed the manipulation of the power state through an episodic recall task, after which, the sense of power scale (Anderson et al., 2012; $\alpha = 0.89$) was used as a power manipulation check. Then, participants were presented with two pairs of images (one is the simple shape, and the other is the specific product mug; see **Appendix D**) to test their preference for angular shape over rounded shape.

Results and Discussion

Manipulation Checks

The results of the manipulation checks show that the power state was successfully manipulated, $F_{(2,163)} = 8.30$, $p < 0.001$. Specifically, participants in the high-power condition ($M = 4.50$, $SD = 1.27$) reported more powerful than those in the low-power condition [$M = 3.42$, $SD = 1.50$; $t_{(105)} = 4.04$, $p < 0.001$] and control condition [$M = 3.98$, $SD = 1.33$; $t_{(112)} = 2.08$, $p = 0.040$). The degree of power reported by participants in the low-power condition was significantly lower than that in the control condition [$t_{(109)} = 2.11$, $p = 0.037$).

Preference

An ANOVA yielded a significant influence of power state on shape preference. For the simple shape, participants in the high-power condition ($M = 4.58$, $SD = 1.63$) reported more positive attitudes toward angular shape over rounded shape than those

in the low-power condition [$M = 3.81$, $SD = 1.66$; $t_{(105)} = 2.44$, $p = 0.017$] and control conditions [$M = 3.93$, $SD = 1.76$; $t_{(112)} = 2.04$, $p = 0.044$], $F_{(2,163)} = 3.32$, $p = 0.039$. But the attitude between low-power condition and control condition has no significant difference [$t_{(109)} = 0.38$, $p = 0.703$]. Meanwhile, for the mug, the effect was consistent with the result of the simple figure. Participants in the high-power condition ($M = 4.38$, $SD = 1.86$) reported more positive attitudes toward angular shape over rounded shape than those in the low-power condition [$M = 3.46$, $SD = 1.87$; $t_{(105)} = 2.55$, $p = 0.012$] and control conditions [$M = 3.56$, $SD = 1.97$; $t_{(112)} = 2.29$, $p = 0.024$], $F_{(2,163)} = 3.86$, $p = 0.023$). The attitude between the low-power condition and control condition was not significantly different [$t_{(109)} = 0.27$, $p = 0.790$].

Studies 1 and 2 verified Hypothesis 1—consumers experiencing higher power are more likely to prefer an angular shape over a rounded shape than those experiencing lower power, demonstrated through different research methods, research objects, and experimental materials and improving the external validity of the conclusions. In Studies 3–5, we further tested the proposed mechanism that underlies the observed effect.

STUDY 3: A DIRECT TEST OF PERCEIVED COMPETENCE AND WARMTH

In order to better understand the relationship between power and shape preference, Study 3 mainly aimed to verify the mediating role of perceived competence and warmth, and to exclude alternative explanations of conflict resolution types and need for uniqueness.

Method

Sample

A total of 288 workers (146 females and 140 males; $M_{age} = 34.45$, $SD_{age} = 11.06$) from the Prolific participant pool took part in the online study in exchange for a \$0.3 reward.

Procedure

Just as in Study 1a, participants first rated the extent to which they agreed with the items of sense of power (Anderson et al., 2012), $\alpha = 0.91$. To capture underlying mechanisms, participants were subsequently required to rate three items relating to the question, "When you look at an object, what characteristic do you value more about it?" Each item was rated on a scale from 1 to 7. The items were: (1) warmth vs. competence; (2) sincere vs. capable; (3) friendly vs. intelligent (Fiske et al., 2002). This was done to measure their valuing of competence and warmth ($\alpha = 0.79$). Higher scores indicated higher values of competence. In addition, in order to exclude alternative explanations of the need for uniqueness and conflict resolution types, participants were then asked to fill in the scales of conflict resolution types (Morris et al., 1998; $\alpha = 0.83$) and need for uniqueness (adapted from Tian and Mckenzie, 2001; $\alpha = 0.81$). We then presented participants with an image of a pair of simple shapes (see **Appendix E**) and asked their preference for the angular and rounded shape (the same as Study 1a). Finally, the participants were debriefed and thanked.

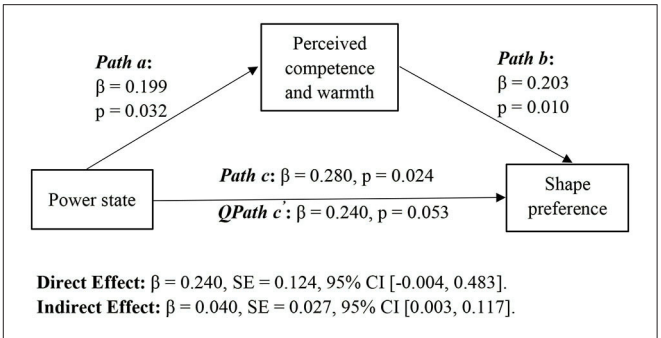


FIGURE 1 | Mediation analyses. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

Results and Discussion

Preference

Results showed participants with a higher sense of power reported more positive attitudes toward angular shape over rounded shape ($r = 0.14$, $p = 0.022$), while controlling for participants' age, sex, and social status, the positive effect was still significant ($\beta = 0.13$, $p = 0.024$).

Mediation Analyses

We performed a regression-based mediation analysis following Hayes' PROCESS macro (Hayes, 2013) to verify the mediating role of perceived competence and warmth. The results in **Figure 1** show valuing competence over warmth mediated the positive relationship between experiencing higher power and the preference for an angular shape (indirect effect = 0.040, 95% bias-corrected CI [0.002, 0.118]). Meanwhile, we also examined the mediating effects of conflict resolution types and need for uniqueness. However, the results revealed that the mediation of conflict resolution types (indirect effect = 0.001, 95% bias-corrected CI [−0.012, 0.034]) and need to be unique (indirect effect = 0.007, 95% bias-corrected CI [−0.128, 0.058]) were not significant. The results of mediation analyses showed consumers experiencing higher power are reliant more on competence over warmth; thus, they are more likely to prefer an angular shape over a rounded shape than consumers experiencing lower power.

STUDY 4: EXPERIMENTAL EVIDENCE OF MEDIATION

Study 3 preliminarily verified the mediating role of perceived competence and warmth through the survey method and indicated that, generally, consumers with high power pay more attention to competence characteristic and are thus more likely to prefer an angular shape, while low-power consumers pay more attention to warmth characteristic and are more likely to prefer a rounded shape. In Study 4, we used experimental manipulation to prime participants' sense of power in order to verify the causal relationship between sense of power and shape preference, and the mediating role of perceived competence and warmth.

Method

Sample and Design

A total of 112 participants took part in this study (57 females and 55 males; $M_{age} = 20.99$, $SD_{age} = 1.79$) and each participant received a ¥3 reward for completing the study. Participants were randomly assigned to one of the two conditions of a single factor 2 (power: high vs. low) between-subjects design.

Procedure

Participants were asked to complete several unrelated tasks. First, we adopted procedures from Dubois et al. (2010) to manipulate participants' power state. Participants were randomly assigned the roles of manager (high-power condition) and employee (low-power condition). Participants in each condition were asked to read and complete a writing task. In the high-power condition, participants received the instructions, "Assume that you are the manager of the design department of an advertising company. You have complete control over coordination tasks, including work progress, the evaluation of subordinates, and the division of rewards. Then, as a leader, what are your feelings and thoughts, and how will you carry out your work and manage your subordinates?" In the low-power condition, participants received the instructions, "Assume that you are the employee of the design department of an advertising company. You have no control over the coordination task, and your performance, salary, and position promotion are judged by the leader. Then, as a subordinate, what are your feelings and thoughts, and how will you treat your work and your leader?" We again used the sense of power scale (Anderson et al., 2012) as a power manipulation check. The scale showed high reliability ($\alpha = 0.88$).

Subsequently, just like Study 3, the participants were asked to rate in three items that contain the corresponding options: When you looked at an object what characteristic do you value more about it (1) 1 = pay more attention to warmth, 7 = pay more attention to competence; (2) 1 = pay more attention to sincere, 7 = pay more attention to capable; (3) 1 = pay more attention to friendly, 7 = pay more attention to intelligent (Fiske et al., 2002), to measure their valuing for competence and warmth (α =0.79). In addition, in order to exclude alternative explanations of the need for uniqueness and conflict resolution types, participants were then asked to fill in the scales of conflict resolution types (Morris et al., 1998; α =0.82) and need for uniqueness (adapted from Tian and Mckenzie, 2001; α =0.82). Then, participants were presented with three pairs of images (see **Appendix F**) to test their preference for angular shape over rounded shape. Finally, the participants were debriefed and thanked.

Results and Discussion

Manipulation Checks

The results of the manipulation check showed that power state was successfully manipulated, $F_{(1,110)}$ =12.65, p =0.001. Specifically, participants in the high-power condition ($M = 4.62$, $SD = 0.97$) reported more powerful than those in the low-power condition ($M = 3.90$, $SD = 1.17$).

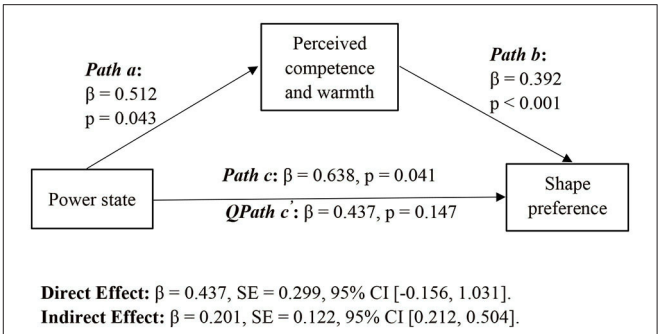


FIGURE 2 | Mediation analyses. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

Preference

Combining participants' preference in the three pairs of images as the dependent variable, an ANOVA yielded a significant influence of power state on shape preference [$F_{(1,110)} = 4.26$, $p = 0.041$]. Specifically, participants in the high-power condition ($M = 4.50$, $SD = 1.62$) reported more positive attitudes toward angular shape over rounded shape than those in the low-power condition ($M =$ 3.86, SD =1.65).

Mediation Analyses

We conducted a mediation analysis with the participants in the high and low power conditions. In accordance with Hayes' (Hayes, 2013) study (model 4), we used a bootstrapping procedure that generated a sample size of 5,000 to assess the regression models. The results of this analysis indicated that participants with higher power valuing more on competence over warmth ($\beta = 0.51$, $t = 2.05$, $p = 0.043$), and the perceived competence has a positive impact on participants' preference toward angular shape over rounded shape ($\beta = 0.39$, $t = 3.50$, $p < 0.001$). Furthermore, the perceived competence and warmth mediated the positive relationship between power state and shape preference (indirect effect = 0.20, 95% bias-corrected CI [0.021, 0.504], direct effect =0.44, 95% bias-corrected CI [−0.156, 1.03]). Meanwhile, we also examined the mediating effects of conflict resolution types and need for uniqueness. However, the results revealed that the mediation of conflict resolution types (indirect effect =0.13, 95% bias-corrected CI [−0.010, 0.419]) and need to be unique (indirect effect =0.05, 95% bias-corrected CI [−0.048, 0.247]) were not significant. The results of mediation analyses showed high power consumers pay more attention on competence over warmth; thus, they are more likely to prefer an angular shape over a rounded shape than low power consumers (**Figure 2**).

STUDY 5: MANIPULATE THE PERCEIVED COMPETENCE AND WARMTH

Study 3 and Study 4 verified the mediation effect through measurement and experiment respectively. However, sometimes, the product itself or its related advertising will also endow the product with certain competence and warmth attributes. Based

on this, when designing a product, the most important subject is how to integrate shape design with product characteristics and product application context. Therefore, in Study 5, we verified the mediating mechanism again by manipulating the perception of competence and warmth in the laboratory. That is, when we manipulated the perceived competence and warmth attributes of marketing scenarios, whether the matching effect between consumers' power state and shape preference will still exist.

Method

Sample and Design

A total of 246 undergraduates completed the study in the lab (134 females and 114 males; $M_{age} = 19.25$, $SD_{age} = 1.12$) and each participant received a ¥3 reward for completing the study. They were randomly assigned to one of the six conditions of a 2 (power states: low vs. high) * 3 (product attributes: competence vs. warmth vs. control) between-subjects design.

Procedure

Firstly, the participants completed the manipulation of the power state just like Study 4, after which, the sense of power scale (Anderson et al., 2012; $\alpha = 0.83$) was used as a power manipulation check.

Subsequently, the participants in each condition were presented with an image of one pair of brand logos (see **Appendix G**, the order of presentation of angular shape and rounded shape was randomized), and informed that these are the pending logos of a certain automobile brand. Participants were asked to read the relevant introduction of the brand and choose the appropriate logo based on their understanding of the brand. In the competence condition, the product introduction emphasizes the performance advantages of the car. In the warmth condition, the product introduction emphasizes the warmth advantages of the car. In the control condition, the product introduction did not specifically emphasize any advantages. Specifically, participants in the competence condition received the information that "A company has recently developed a new type of car, which has strong power performance. The overall design adopts a high-tech and high configuration, with a strong driving force and excellent performance. At present, the company wants to create a simple and memorable brand logo for the car." Participants in the warmth condition received the information that "A company has recently developed a new type of car, which is mainly aimed at the family car market, focusing on warmth and comfort. The design process fully reflects care and humanization. At present, the company wants to create a simple and memorable brand logo for the car." Participants in the control condition received the instructions, "A company has recently developed a new type of car; at present, the company wants to create a simple and memorable brand logo for the car." Then, the participants were asked to express their opinion on the suitability of the following two brand logos based on their understanding of the car (1 = A is very suitable, 7 = B is very suitable). Finally, the participants were debriefed and thanked.

Results and Discussion

Manipulation Checks

An ANOVA performed on participants' sense of power yielded only a main effect of power such that the participants in the high-power condition reported more powerful ($M = 4.15$, $SD = 1.49$) than those in the low-power condition ($M = 3.72$, $SD = 1.28$), $F_{(1,246)} = 5.85$, $p = 0.016$. None of the other effects were significant ($p > 0.28$).

Preference

With participants' shape preference as the dependent variable, the ANOVA revealed significant main effects of power states [$F_{(1,242)} = 6.39$; $p = 0.012$] and product attributes [$F_{(2,242)} = 34.81$; $p < 0.001$). There was also a significant interaction between power states and product attributes [$F_{(2,242)} = 3.09$, $p = 0.047$). Specifically, in the control condition of product attributes, high-power participants reported higher preference ($M = 4.71$, $SD = 1.60$) toward angular shape over rounded shape than low-power participants ($M = 3.61$, $SD = 1.43$), $t_{(81)} = 3.32$, $p = 0.001$. Most importantly, compared with control conditions, in the competence condition, both high-power participants ($M = 4.85$, $SD = 1.78$) and low-power participants ($M = 4.78$, $SD = 1.29$) preferred an angular shape over rounded shape, and the difference between them was not significant [$t_{(80)} = 0.21$, $p = 0.832$]. On the contrary, in the warmth condition, both high-power participants ($M = 3.07$, $SD = 1.23$) and low-power participants ($M = 2.86$, $SD = 1.26$) preferred rounded shape over angular shape, and the difference between them was not significant [$t_{(81)} = 0.79$, $p = 0.432$].

The results of the experiment showed when the products' competence and warmth attributes were not manipulated (control condition), compared with low-power consumers, high-power consumers prefer an angular shape over a rounded shape, consistent with Hypothesis 1. When the products' competence and warmth attributes were manipulated, the effect of power states on shape preference disappeared, again verifying the mediating effect of competence and warmth (Hypothesis 2).

GENERAL DISCUSSION

In the current research, we conducted four studies and provided supportive evidence for the hypothesis that consumers in different power states respond differently to angular and rounded shapes. Study 1 provided preliminary evidence that consumers with a higher power state prefer an angular (vs. rounded) shape design than those with a lower power state. In Study 2, the causal relationship between them was further verified by an experimental method. Study 3 and Study 4 further verified the underlying mechanism of perceived competence and warmth; specifically, high-power consumers responded to the angular shape more positively because consumers experiencing higher power are valuated more on competence over warmth, thus they are more likely to prefer an angular shape over a rounded shape than consumers experiencing lower power. Study 5 provided further evidence of the underlying mechanism. That is, when we manipulated the perceived competence and warmth of the product attribute, the shape of the product or brand logo is

integrated with its attributes; then, at this time, the matching effect between power state and shape preference disappears.

Theoretical and Practical Implications

Our research contributes to the literature in several ways. First, the conclusions of this study expand the research of power state in the field of sensory marketing. Existing studies on power state mainly focus on persuasive messages (Dubois et al., 2016), compensatory consumption (Dubois et al., 2012), spending preferences (Rucker et al., 2011), and price fairness perceptions (Jin et al., 2014), but are not involved in the field of sensory marketing. However, sensory marketing has been a rapidly growing field in recent years (Krishna, 2012; Bajaj and Bond, 2018; Mookerjee et al., 2021) and has an important position in consumer behavior research. The current research, therefore, explores the preference effect of power state on visual shapes that have an important influence in sensory marketing, and proposes the matching effect between the two driven by the psychological propensities of power, which provides a new research idea for the power state and enriches the literature on power in the field of consumer behavior.

Second, as crucial visual elements, angular and rounded shapes have been widely studied by scholars and marketers (Bajaj and Bond, 2018; Vallen et al., 2018; Yang and Chen, 2019). However, the existing research on shape mainly focuses on the marketing effect brought about by shape design. For example, Yang and Chen (2019) demonstrated that in the context of precise numbers (vs. round numbers), the presentation of angular brand logos led to higher consumer evaluation than the presentation of rounded brand logos. However, studies on the factors that affect consumers' preference for different shapes are relatively scarce. Therefore, the current research attempts, from the perspective of psychology, to explore the relationship between power state and consumers' shape preference, enriching the literature of shape design and providing a new perspective for the study of sensory marketing.

Finally, a new underlying mechanism is introduced to explain the relationship between power and shape preference. The existing mechanism of studies on angular and rounded shapes focused on conflict resolution (Morris et al., 1998) and need for uniqueness (Tian and Mckenzie, 2001), which cannot adequately explain our current theory. In this paper, we introduce a new mediation mechanism, i.e., perceived competence and warmth. As two fundamental dimensions of social judgments, marketing scholars have widely studied perceived competence and warmth in recent years, and many important research conclusions have been obtained (Dubois et al., 2016). However, the existing research on competence and warmth mainly focuses on the perception and judgment of people or organizations but does not involve products or goods. Thus, this study initiates a new research field for scaling competence and warmth.

This study has several practical implications for advertisers and marketing professionals. First, our conclusions have important marketing implications for targeting different groups of customers. Specifically, the results show that high-power consumers prefer angular shape design, while low-power consumers prefer rounded shape design. Therefore, in practice, marketers can perform market segmentation based on variables correlated with the sense of power, such as socioeconomic status to design or recommend different products for consumers with different power states. Second, due to the scenario and maneuverability of power state, in the actual marketing strategy, marketers can provide advertising and environmental cues that shift consumers' current state of power according to their own product and brand characteristics, to achieve consistency in consumer power state and product characteristics, improving consumers' purchase intention and reducing their price sensitivity. Finally, graphic shapes are a very important part of marketing. This study points out that different shapes have different characteristics. In particular, the angular shape reflects more on competence, while rounded shape reflects more on warmth. According to this conclusion, businesses can design their brands and product shapes according to the marketing connotations they pursue to optimize the "self-expression" of products and obtain better consumer evaluations.

Research Limitations and Future Research Directions

First, the influence of power state on consumers' thoughts and behaviors is mostly performed from the perspective of psychological propensities and psychological needs. This paper is based on the psychological propensities of power to explore the shape preferences of different power states. However, at present, the relevant research is based on a single perspective of psychological propensities or psychological needs. Through theoretical retrieval and experimental testing, we find that psychological propensities and psychological needs may jointly affect consumption behavior and their impact on individuals' consumption behavior may be different. Therefore, future research can follow this line of thinking to explore the different effects of power's psychological propensities and psychological needs impact on the same consumption phenomenon, which may lead to new findings.

Second, this paper explores the influence of power states on the preference for angular and rounded shapes, enriching the literature in the field of visual design. In the field of visual sense, in addition to shape, there are other visual elements. For example, Bajaj and Bond (2018) explored the relationship between the symmetry of graphics and brand personality. In future research, we will further explore the effect of regular and irregular graphics and complex and simple graphics. In addition to visual elements, the influence of color on consumer behavior has also gradually developed in recent years (Lee et al., 2016; Hagtvedt and Brasel, 2017), and new exploration in this field can also be performed in the future.

DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author.

ETHICS STATEMENT

Ethical review and approval were not required for this type of study on human participants in accordance with

the local legislation and institutional requirements. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

FY designed the study, analyzed and interpreted the data, conceived and drafted the manuscript. XJ participated in the interpretation of results and edited the manuscript. BW provided important reference opinions for research design, and critically revised the manuscript. TZ conducted data collection and conducted writing instruction. TM co-designed the study, and participated in the data collection. All authors contributed to the article and approved the submitted version.

FUNDING

The current study was supported by the National Natural Science Foundation of China (Grant 71872070, 71902069, 71902148), Social Science Foundation project of Jilin Province (Grant 2020C060), China Postdoctoral Science Foundation (Grant 2019M660243), and Sichuan University (Grant 2020CXQ23).

ACKNOWLEDGMENTS

We express our sincere gratitude to the participants who are involved in the present research and to the teachers who help with the data collection. Moreover, we much appreciate the editor's and the reviewers' valuable time and fine-grained suggestions in terms of revisions to the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2021.615647/full#supplementary-material

REFERENCES

Aaker, J., Vohs, K. D., and Mogilner, C. (2010). Nonprofits are seen as warm and for-profits as competent: Firm stereotypes matter. *J. Consum. Res.*. 37, 224–237. doi: 10.1086/651566

Anderson, C., John, O. P., and Keltner, D. (2012). The personal sense of power. *J. Pers.*. 80, 313–344. doi: 10.1111/j.1467-6494.2011.00734.x

Arnheim, R. (1974). Art and Visual Perception: A Psychology of the Creative Eye. Berkeley, CA: University of California Press. doi: 10.1525/9780520351271

Bajaj, A., and Bond, S. D. (2018). Beyond beauty: design symmetry and brand personality. *J. Consum. Psychol.* 28, 77–98. doi: 10.1002/jcpy.1009

Bakan, D. (1966). Test of significance in psychological research. *Psychol. Bull.* 66, 423–437. doi: 10.1037/h0020412

Blader, S. L., and Chen, Y. (2012). Differentiating the effects of status and power: a justice perspective. *J. Pers. Soc. Psychol.* 102, 994–1014. doi: 10.1037/a0026651

Blader, S. L., Shirako, A., and Chen, Y. R. (2016). Looking out from the top: differential effects of status and power on perspective taking. *Pers. Soc. Psychol. Bull.* 42, 723–737. doi: 10.1177/0146167216636628

Brinol, P., Petty, R. E., Valle, C., Rucker, D. D., and Becerra, A. (2007). The effects of message recipients' power before and after persuasion: a self-validation analysis. *J. Pers. Soc. Psychol.* 93, 1040–1053. doi: 10.1037/0022-3514.93.6.1040

Dubois, D., Rucker, D. D., and Galinsky, A. D. (2010). The accentuation bias: money literally looms larger (and sometimes smaller) to the powerless. *Soc. Psychol. Pers. Sci.* 1, 199–205. doi: 10.1177/1948550610365170

Dubois, D., Rucker, D. D., and Galinsky, A. D. (2012). Super size me: product size as a signal of status. *J. Consum. Res.* 38, 1047–1062. doi: 10.1086/661890

Dubois, D., Rucker, D. D., and Galinsky, A. D. (2016). Dynamics of communicator and audience power: the persuasiveness of competence versus warmth. *J. Consum. Res.* 43, 68–85. doi: 10.1093/jcr/ucw006

Fiske, S. T., Cuddy, A. J., and Glick, P. (2007). Universal dimensions of social cognition: warmth and competence. *Trends Cogn. Sci.* 11, 77–83. doi: 10.1016/j.tics.2006.11.005

Fiske, S. T., Cuddy, A. J., Glick, P., and Xu, J. (2002). A model of (often mixed) stereotype content: competence and warmth respectively follow from perceived status and competition. *J. Pers. Soc. Psychol.* 82, 878–902. doi: 10.1037/0022-3514.82.6.878

Galinsky, A. D., Gruenfeld, D. H., and Magee, J. C. (2003). From power to action. *J. Pers. Soc. Psychol.* 85, 453–466. doi: 10.1037/0022-3514.85.3.453

Galinsky, A. D., Magee, J. C., Gruenfeld, D. H., Whitson, J. A., and Lilenquist, K. A. (2008). Power reduces the press of the situation: Implications for creativity, conformity, and dissonance. *J. Pers. Soc. Psychol.* 95, 1450–1466. doi: 10.1037/a0012633

Galinsky, A. D., Rucker, D. D., and Magee, J. C. (2015). "Power: Past findings, present considerations, and future directions," in *APA handbook of personality and social psychology, Vol. 3: Interpersonal Relationships*, eds M. Mikulincer, and P. R. Shaver (Washington, DC: American Psychological Association), 421–460. doi: 10.1037/14344-016

Garbinsky, E. N., Klesse, A., and Aaker, J. (2014). Money in the bank: feeling powerful increases saving. *J. Consum. Res.* 41, 610–623. doi: 10.1086/676965

Hagtvedt, H., and Brasel, S. A. (2017). Color saturation increases perceived product size. *J. Consum. Res.* 44, 396–413. doi: 10.1093/jcr/ucx039

Hayes, A. F. (2013). Introduction to mediation, moderation, and conditional process analysis. *J. Educ. Meas.* 51, 335–337. doi: 10.1111/jedm.12050

Jiang, Y., Gorn, G. J., Galli, M., and Chattopadhyay, A. (2016). Does your company have the right logo? How and why circular- and angular-logo shapes influence brand attribute judgments. *J. Consum. Res.* 42, 709–726. doi: 10.1093/jcr/ucv049

Jin, L., He, Y., and Zhang, Y. (2014). How power states influence consumers' perceptions of price unfairness. *J. Consum. Res.* 40, 818–833. doi: 10.1086/673193

Jin, L., and Huang, Y. (2018). How power states influence the persuasiveness of top-dog versus underdog appeals. *J. Consum. Psychol.* 29, 243–261. doi: 10.1002/jcpy.1069

Jin, L., Zou, D., He, Y., and Xu, Q. (2011). A sense of power and uniqueness seeking. *Adv. Consum. Res.* 9:376

Keltner, D., Gruenfeld, D. H., and Anderson, C. (2003). Power, approach, and inhibition. *Psychol. Rev.* 110, 265–284. doi: 10.1037/0033-295X.110.2.265

Kim, S., and McGill, A. L. (2011). Gaming with Mr. Slot or gaming the slot machine? Power, anthropomorphism, and risk perception. *J. Consum. Res.* 38, 94–107. doi: 10.1086/658148

Krishna, A. (2012). An integrative review of sensory marketing: engaging the senses to affect perception, judgment and behavior. *J. Consum. Psychol.* 22, 332–351. doi: 10.1016/j.jcps.2011.08.003

Lee, H., Fujita, K., Deng, X., and Unnava, H. R. (2016). The role of temporal distance on the color of future-directed imagery: a construal-level perspective. *J. Consum. Res.* 43, 707–725. doi: 10.1093/jcr/ucw051

Lieven, T., Grohmann, B., Herrmann, A., Landwehr, J. R., and Van Tilburg, M. (2015). The effect of brand design on brand gender perceptions and brand preference. *Eur. J. Market.* 49, 146–169. doi: 10.1108/EJM-08-2012-0456

Lundholm, H. (1921). The affective tone of lines: experimental researches. *Psychol. Rev.* 28, 43–60. doi: 10.1037/h0072647

Magee, J. C., and Galinsky, A. D. (2008). Social hierarchy: the self-reinforcing nature of power and status. *Acad. Manage. Ann.* 2, 351–398. doi: 10.5465/19416520802211628

Mookerjee, S. S., Cornil, Y., and Hoegg, J. J. (2021). Express: from waste to taste: how "ugly" labels can increase purchase of unattractive produce. *J. Market.* 85, 62–77. doi: 10.1177/0022242920988656

Morris, M. W., Williams, K. Y., Leung, K., Larrick, R. P., Mendoza, M. T., Bhatnagar, D., et al. (1998). Conflict management style: accounting for cross-national differences. *J. Int. Bus. Stud.* 29, 729–747. doi: 10.1057/palgrave.jibs.8490050

Poffenberger, A. T., and Barrows, B. E. (1924). The feeling value of lines. *J. Appl. Psychol.*. 8, 187–205. doi: 10.1037/h0073513

Roux, C., Goldsmith, K., and Bonezzi, A. (2012). *On the Consequences of a Scarcity Mindset: Why Thoughts of Having Less Can Lead to Taking (and Giving) More*. Rochester, NY: Social Science Electronic Publishing.

Rucker, D. D., Dubois, D., and Galinksy, A. D. (2011). Generous paupers and stingy princes: power drives consumer spending on self- versus others. *J. Consum. Res.* 37, 1015–1029. doi: 10.1086/657162

Rucker, D. D., and Galinsky, A. D. (2008). Desire to acquire: powerlessness and compensatory consumption. *J. Consum. Res.* 35, 257–267. doi: 10.1086/588569

Rucker, D. D., and Galinsky, A. D. (2009). Conspicuous consumption versus utilitarian ideals: how different levels of power shape consumer behavior. *J. Exp. Soc. Psychol.* 45, 549–555. doi: 10.1016/j.jesp.2009.01.005

Rucker, D. D., Galinsky, A. D., and Dubois, D. (2012). Power and consumer behavior: how power shapes who and what consumers value. *J. Consum. Psychol.* 22, 352–368. doi: 10.1016/j.jcps.2011.06.001

Rucker, D. D., Galinsky, A. D., and Magee, J. C. (2018). The agentic–communal model of advantage and disadvantage: How inequality produces similarities in the psychology of power, social class, gender, and race. *Adv. Exp. Soc. Psychol.* 58, 71–125. doi: 10.1016/bs.aesp.2018.04.001

Sirgy, M. J. (1982). Self-concept in consumer behavior: a critical review. *J. Consum. Res.* 9, 287–300. doi: 10.1086/208924

Sirgy, M. J., Grewal, D., and Mangleburg, T. F. (2004). Retail environment, self-congruity, and retail patronage: an integrative model and a research agenda. *J. Bus. Res.* 49, 127–138. doi: 10.1016/S0148-2963(99)00009-0

Tian, K., and Mckenzie, K. (2001). The long-term predictive validity of the consumers' needs for uniqueness scale. *J. Cons. Psychol.* 10, 171–193. doi: 10.1207/s15327663jcp1003_5

Vallen, B., Sridhar, K., Rubin, D., Ilyuk, V., Block, L. G., and Argo, J. J. (2018). Shape- and trait-congruency: Using appearance-based cues as a basis for product recommendations. *J. Consum. Psychol.* 29, 271–284. doi: 10.1002/jcpy.1065

Yang, C., and Chen, Z. X. (2019). Does number have shape? the matching effect of digital information accuracy and brand logo shape. *Acta Psychol. Sinica* 51, 841–856. doi: 10.3724/SP.J.1041.2019.00841

Zhang, Y., Feick, L., and Price, L. J. (2006). The impact of self-construal on aesthetic preference for angular versus rounded shapes. *Pers. Soc. Psychol. Bull.* 32, 794–805. doi: 10.1177/0146167206286626

Zhu, R., and Argo, J. J. (2013). Exploring the impact of various shaped seating arrangements on persuasion. *J. Consum. Res.* 40, 336–349. doi: 10.1086/670392

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Yao, Jin, Wu, Zhao and Ma. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Psychology

BRIEF RESEARCH REPORT
published: 05 January 2022
doi: 10.3389/fpsyg.2021.786507

Hong Kong Women Project a Larger Body When Speaking to Attractive Men

Albert Lee[1]* *and Eva Ng*[2]

[1] Department of Linguistics and Modern Language Studies, The Education University of Hong Kong, Tai Po, Hong Kong SAR, China, [2] Department of Linguistics, The University of Hong Kong, Pokfulam, Hong Kong SAR, China

In this pilot study we investigated the vocal strategies of Cantonese women when addressing an attractive vs. unattractive male. We recruited 19 young female native speakers of Hong Kong Cantonese who completed an attractiveness rating task, followed by a speech production task where they were presented a subset of the same faces. By comparing the rating results and corresponding acoustic data of the facial stimuli, we found that when young Cantonese women spoke to an attractive male, they were less breathy, lower in fundamental frequency, and with denser formants, all of which are considered to project a larger body. Participants who were more satisfied with their own height used these vocal strategies more actively. These results are discussed in terms of the body size projection principle.

Keywords: body size projection, Cantonese, sexual selection, sociophonetics, vocal attractiveness

OPEN ACCESS

Edited by:
Oliver Niebuhr,
University of Southern Denmark, Denmark

Reviewed by:
Kara Hawthorne,
Gallaudet University, United States
Katia Nemr,
University of São Paulo, Brazil

***Correspondence:**
Albert Lee
albertlee@eduhk.hk

Specialty section:
This article was submitted to Language Sciences, a section of the journal Frontiers in Psychology

Received: *30 September 2021*
Accepted: *13 December 2021*
Published: *05 January 2022*

Citation:
Lee A and Ng E (2022) Hong Kong Women Project a Larger Body When Speaking to Attractive Men. Front. Psychol. 12:786507. doi: 10.3389/fpsyg.2021.786507

INTRODUCTION

Having an attractive voice is useful because listeners tend to associate it with an attractive face (Hughes and Miller, 2016), a likeable personality (Zuckerman and Driver, 1989), and assign it higher health ratings (Albert et al., 2021). It has been reported that physical attractiveness leads to advantages in situations such as dating (Berscheid et al., 1971), job applications (Watkins and Johnston, 2000), promotion (Chung and Leung, 1988), elections (Jäckle et al., 2020), and is associated with more social support (Sarason et al., 1985). While one's physical appearance cannot be easily altered at least in the short run, adjusting their own voice is an immediately possible alternative. Therefore, a good understanding of vocal attractiveness is of practical, social, and theoretical importance. So far, researchers have identified the characteristics of an attractive voice in perception experiments (e.g., by rating voice stimulus), but whether speakers choose to speak in the same preferred voice is an open question. This study approached this lesser-studied aspect of vocal attractiveness by studying how Cantonese women from Hong Kong choose their vocal strategies when addressing attractive vs. unattractive men.

Averageness vs. Body Size Projection

Two seemingly competing hypotheses seek to account for the phonetics of an attractive voice, namely the averageness hypothesis and body size projection. The former stems from the 'averaging attractiveness phenomenon' and argues that voices similar to the population average are considered more attractive (see review in Belin, 2021). From an evolutionary point of view,

the average voice may signal good genes as it has withstood evolution and adaptive changes to become the norm of the population, much like the average face appearing to signal high fitness (Langlois and Roggman, 1990; Thornhill and Gangestad, 1999). From a perceptual perspective, the average voice may be easier to process as it resembles a central voice prototype based on which voice identities are encoded, as is the case for face (Winkielman et al., 2006).

Meanwhile, the body size projection principle (Morton, 1977) contends that animals use their voice to project different body sizes to serve different communicative functions (e.g., small projected body to show appeasement, large to express hostility). Extending this principle, subsequently Xu et al. (2013) found that an attractive male voice to female English listeners was one that sounded large, vice versa for a female voice to male listeners. However, this does not mean that, for example, perceived attractiveness would monotonously increase with a smaller projected female body – extremely high fundamental frequency (f_o henceforth, i.e., the acoustic correlate of pitch) was not judged as the most attractive in Xu et al. (2013), possibly because the very small projected body started to sound more child-like than attractive. All in all, it appears that an attractive voice is one that resembles the population average, with specific projected body sizes (larger for male speakers, smaller for female speakers) adding enhancing effects, provided they do not deviate too much from the average.

Acoustic Correlates of Body Size

In general, there is an inverse relationship between body size and f_o (Morton, 1977) as well as formant dispersion (Fitch, 1997). f_o is the frequency at which membranes (the vocal folds in the case of humans) vibrate (see Lee and Mok, 2021 for a recent review), and is determined by body size – "(t)he larger the animal, the lower the sound frequency it can produce" (Morton, 1977, p. 864). Formant dispersion, or the averaged difference between successive formant frequencies, reflects one's vocal tract length (Fitch, 1997), and in turn body size. The shorter the vocal tract, the further apart the speaker's formants. As for f_o *range*, the use of a larger f_o range is associated with both cooperativeness (*cf.* "the effort code," Gussenhoven, 2016) and happiness (Xu et al., 2013), in turn likely a smaller body which signals less threat. In terms of voice quality, breathy voice (acoustically manifested in spectral parameters such as "H1–A1" and "H1–A3") is argued to be acoustically more similar to pure tone compared with voice qualities such as modal voice (Xu et al., 2013), and signals a small body according to Morton (1977). Conversely, creaky voice (main acoustic correlates: "jitter" and "shimmer") is typically argued to be associated with masculinity and authority (see Yuasa, 2010 for a review).

Cross-Linguistic Variation in Preferences in Voices

The acoustic correlates of an attractive voice have been extensively studied in recent years. To male English listeners, an attractive female voice is high in f_o, breathy, and with wide formant dispersion, all of which signal a small body; to female English listeners, an attractive male voice (i) is low in f_o and (ii) has narrow formant dispersion, both signaling a large body, but (iii) is also breathy, signaling a smaller body (Xu et al., 2013), presumably to neutralize some of the hostility accompanying the large projected body.

It has been reported that the creaky voice is increasingly used by American female speakers in recent years (Yuasa, 2010). Although this seems to deviate from the body-size projection principle, as creakiness is considered to be associated with a large body, there is also evidence that the use of creaky voice by American women is considered less attractive than a normal speaking voice (Anderson et al., 2014). Therefore, it seems to suggest instead that speakers' vocal strategies do not necessarily have to align with what the opposite sex considers attractive.

Apart from Xu et al. (2013), comparable perception studies on non-Western populations include Japanese (Xu et al., 2017) and Mandarin (Xu and Lee, 2018), which demonstrated cross-linguistic variations in the acoustic cues to an attractive voice. These studies found that while the general principles of body size projection in accounting for patterns in voice preferences appeared to hold, there were also language-specific deviations. For example, in Mandarin and Japanese, a narrow f_o range (which signals a larger body) was found to be unattractive to both male and female listeners alike.

Although the perception of vocal attractiveness in western societies is relatively well understood, there is much less production data available, let alone from non-Western populations. To the best of our knowledge, to date there is no production study on vocal attractiveness in Cantonese. This study serves to fill this gap. While perception studies are useful for identifying the effect of individual acoustic cues, production data are essential as they show how these cues interact in everyday speech. In addition, production data can shed light on individual variability, which is increasingly important with the emergence of statistical tools capturing speakers as a random factor.

Hypotheses

Based on the studies reviewed above, we expected that female Cantonese speakers would use vocal strategies to signal a small body (Hypotheses 3 and 4) when addressing an attractive male, but they might also be creaky (related to Hypotheses 1, 2, 5, and 6) like their American counterparts. The seemingly arbitrary prediction of creakiness is based on two reasons: (i) we tested well-educated young women in Hong Kong where the influence of western (including American) culture is prevalent (Louie, 2010), and (ii) the effect of voice quality in neutralizing one's projected body size was also observed in female English listeners' preferences in a male voice (Xu et al., 2013). Therefore, in this study we tested the following hypotheses (see **Table 1**):

Hypothesis 1 and 2 are related to the use of breathy voice. As the decrease in energy at higher frequencies from the first harmonic (or H1) is the greatest for breathy voice and the least for creaky voice (see review in Gordon and Ladefoged, 2001), we expected to see decreased H1–A1 (where A1 stands for amplitude of the first formant) and H1–A3 in the attractive face condition (i.e., less breathy as we are also hypothesizing increased creakiness in Hypotheses 5 and 6). Here we included

TABLE 1 | *Working hypotheses* (prediction for the attractive facial stimulus condition).

	Perceptual property	Acoustic correlate	Hypothesis	Projected body
Hypothesis 1	Breathiness	H1-A1	Decrease	Large
Hypothesis 2		H1-A3	Decrease	Large
Hypothesis 3	Apparent vocal tract length	Formant disp.	Increase	Small
Hypothesis 4	Pitch	Median f_0	Increase	Small
Hypothesis 5	Creakiness	Jitter	Increase	Large
Hypothesis 6		Shimmer	Increase	Large

multiple spectral parameters (i.e., both H1–A1 and H1–A3) to ensure reliability of our results (*cf.* Kreiman et al., 2007). Formant dispersion (Hypothesis 3) is inversely related to vocal tract length, thus in the attractive face condition we expect to see more dispersed formants that project a shorter vocal tract, in turn a smaller body. Hypothesis 4 is based on the assumption that Cantonese women would project a smaller body with higher median f_0 when addressing an attractive man. Finally, while there are different types of creaky voice (Redi and Shattuck-Hufnagel, 2001), each with its own acoustic properties, as working hypotheses (Hypothesis 5 and 6) we hypothesized that Cantonese women would exhibit more cycle-to-cycle variability in the attractive face condition, thus increased jitter and shimmer (i.e., more creakiness).

METHODS

Participants

Nineteen women participated in this study. They were all recruited in Hong Kong, speaking Cantonese as their first language, and university-educated (either then-current students or recently graduated). They aged between 19 and 25, and self-declared as heterosexual. All of them also spoke English and Mandarin as second languages. Their mean height was 159.4 cm (SD ± 4.4). Participation was voluntary and no one received any monetary remuneration. No one reported any (history of) speech and hearing impairment.

Warm-Up Task

This study comprised three tasks: warm-up, facial attractiveness rating, and speech production task. All tasks were completed in the same session in a quiet room on university campus. Participants were recorded using a Logitech H340 microphone at a sampling rate of 44.1 kHz.

During the warm-up session, participants were asked to say the semantically neutral utterance 你好啊,你讀咩科嫁？ "Hello. What is your major?" three times without being presented any visual stimuli. The purpose of this task was to familiarize the participants with main production task, which will be described below.

Facial Attractiveness Rating

Fifty different male facial stimuli were used. We only included faces of East Asian ethnicities as their features are more familiar to our participants (*cf.* Coetzee et al., 2014). Forty of the faces were relatively attractive Asian male faces (celebrities and otherwise). The remaining stimuli were relatively less attractive male faces (again including celebrities). The images of male celebrities were those from Hong Kong, Korea, Japan, and Mainland China. All stimuli were publicly available images obtained from the Internet.

Participants were asked to rate the attractiveness of these 50 faces on a 1 ∼ 10 scale (10 = most attractive) and write down their response on an answer sheet. They were told to base their ratings purely on how much they were attracted to each face, and to ignore any past knowledge of the respective males or experience they might have with people of similar appearances. Stimuli were presented in a randomized order in Microsoft Powerpoint slides.

Production Task

Based on the ratings from above, for each participant the five most attractive and five least attractive faces were used as target stimuli in a subsequent production task. In the event of faces with the same rating, those that were presented later were chosen. Each face was presented three times on separate occasions in random order. Participants were instructed to imagine themselves in a classroom setting, and that the male face was of a classmate sitting next to them. Participants were then to ask the male classmate 你好啊,你讀咩科嫁？ "Hello. What is your major?" From each participant, 30 utterances were recorded. Recordings were subsequently analyzed using ProsodyPro (Xu, 2013 ver. 5.7.2), which allows manual checking of vocal pulses and automatically extracts numerous acoustic measurements, as will be presented below.

Post hoc Questionnaire

Preliminary data analysis revealed a bimodal distribution which was seemingly related to participants' height. Specifically, we seemed to observe that taller participants seemed to behave in the opposite direction from the rest. To verify this, we sent out a questionnaire to gather information on participants' height and how satisfied they were about it. There were four questions in the questionnaire: (1) "How tall are you?," (2) "On a scale of 1 to 10, how satisfied are you about your own height?," (3) "If you are not satisfied, how much taller/shorter would you like to be (in centimeters)?," and (4) "What are you doing to address your unsatisfactory height (e.g., wearing high heels)?" All participants bar one responded (i.e., N = 18). Based on their response, participants were then classified in terms of how satisfied they were about their height, namely (H)ighly satisfied, (M)oderately

satisfied, and (L)east satisfied. There were six participants in each category. The correlation between participants' height and their satisfaction with their own height was nearly but not significant, $r_s = 0.446$, $N = 18$, $p = 0.063$.

RESULTS

We set out to test six hypotheses (see **Table 1**) to examine whether Cantonese women project a small body when addressing an attractive male. Results are shown in **Figure 1**, where attractive (A) and unattractive (U) facial stimuli are compared (coral and turquoise, respectively) for each acoustic correlate of vocal attractiveness. The *X*-axis of **Figure 1** represents how much speakers were satisfied with their own body height (converted into the three categories H, M, L, with H being the most satisfied). See also **Supplementary Figure 1** for corresponding boxplots with height satisfaction contrasts collapsed.

For voice quality, H1–A1 was higher for unattractive stimuli, indicating more use of breathiness when participants spoke to an unattractive face; the same was true for H1–A3. Cantonese women also appeared to lengthen their vocal tract with denser formants in the Attractive condition, thus projecting a larger body. Similarly, participants' median f_0 was lower in the

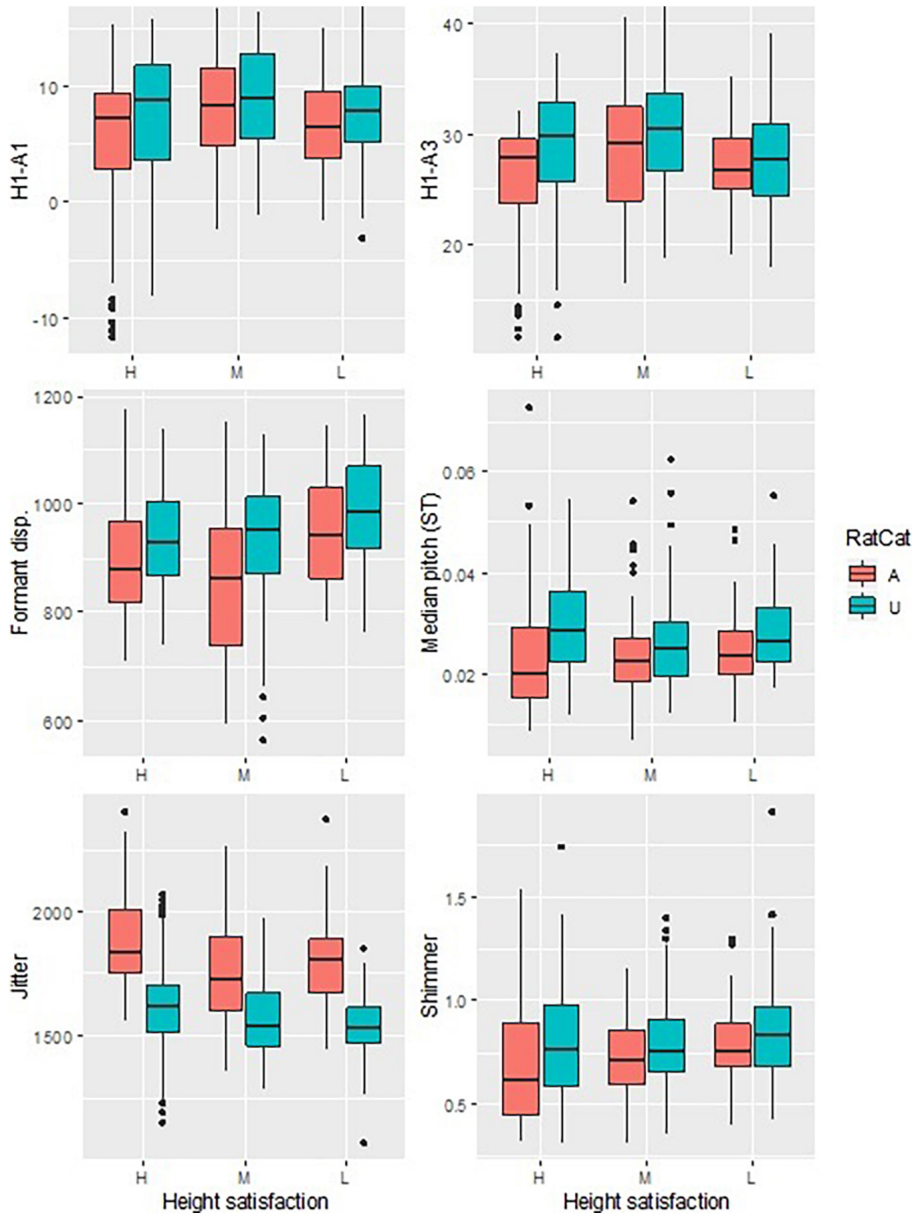


FIGURE 1 | Boxplots comparing acoustic correlates of projected voices in attractive (A) vs. unattractive (U) facial stimulus conditions. H, M, L stand for highly, moderately, least satisfied with speakers' own height.

TABLE 2 | Model summaries for different acoustic correlates.

		Fixed effects					Random effect
		Est.	*SE*	*df*	*t*	*p*	SD Speaker
H1–A1	(Intercept)	8.218	0.930	18.056	8.840	**<0.001**	3.944
	Attract.	−0.417	0.119	17.124	−3.499	**0.003**	0.317
	Attract.:DesChg	0.055	0.021	17.542	2.578	**0.019**	
H1–A3	(Intercept)	29.119	1.218	17.946	23.904	**<0.001**	5.218
	Attract.	−0.444	0.135	19.920	−3.277	**0.004**	0.403
	Attract.: DesChg	0.055	0.022	17.519	2.453	**0.025**	
F1–F3	(Intercept)	968.015	21.398	17.765	45.238	**<0.001**	89.719
	Attract.	−8.909	2.179	15.555	−4.088	**0.001**	8.377
Median f_o	(Intercept)	0.029	0.001	18.489	22.055	**<0.001**	0.005
	Attract.	−0.001	<0.001	15.992	−4.265	**0.001**	0.001
Jitter	(Intercept)	1527.525	36.888	17.853	41.410	**<0.001**	156.64
	Attract.	44.184	5.347	17.882	8.263	**<0.001**	15.78
	Attract.:DesChg	−1.970	0.903	16.526	−2.180	**0.044**	
Shimm.	(Intercept)	0.823	0.036	22.020	23.012	**<0.001**	0.144
	Attract.	−0.015	0.003	542.600	−4.692	**<0.001**	
	Attract.: DesChg	0.001	0.001	480.400	2.011	**0.045**	

"Attract." stands for facial attractiveness rating (1–10, 10 = the most attractive). "DesChg" stands for desired change in height (in cm). Significant fixed effects are in bold.

Attractive condition. In terms of creakiness, participants showed higher jitter but lower shimmer in the Attractive condition.

Initial exploratory data analysis (based on visual inspection of **Supplementary Figure 2**) revealed substantial individual variability in vocal strategies. Therefore, for each acoustical parameter in **Figure 1**, we fitted a linear mixed effects model using the *lmerTest* package in R (Kuznetsova et al., 2017, ver. 3.1-3) to model by-speaker variations. Model summaries are shown in **Table 2**. All models contained the continuous predictor of Attractiveness (rating of male facial stimuli). In some models, we also included the interaction between Attractiveness and desired change in height (see Question 3 in §2.5), which appeared to be a good heuristic of the individual variation. No other manipulation of the data was performed. All models included by-speaker random intercepts; most also included by-speaker random slope for Attractiveness (except for shimmer, in which model including the random slope for Attractiveness would lead to non-convergence).

Table 2 shows that there was a significant main effect of Attractiveness ($p < 0.005$ for all cases) on all acoustical correlates of vocal attractiveness analyzed. This indicates that, after taking into account by-speaker variation, in general an attractive male face elicited significantly less breathiness (lower H1–A1 and H1–A3), longer vocal tract (denser formant dispersion), lower median f_o, less regular cycle-to-cycle variation in f_o (higher jitter) but more regular cycle-to-cycle variation in amplitude (lower shimmer). In addition, there was a significant interaction between Attractiveness and desired change in height in all voice quality-related measurements. It can be seen in **Figure 1** and **Table 3** that the contrast between attractive and unattractive faces in terms of acoustic correlates were bigger for speakers who were satisfied with their own height (in bold in **Table 3**). The opposite parameter estimates of the main effect of Attractiveness and the interaction term in **Table 2** may be understood from the fact that those desiring the largest change in height are the least satisfied, and are using vocal strategies *less* differently to address attractive vs. unattractive males (*cf*. **Table 3**).

DISCUSSION

Summary of Findings

This study explored how Cantonese women projected their voice when speaking to an attractive vs. unattractive face. Results showed that, in the attractive face condition, most acoustic cues pointed to a larger body (except Hypothesis 6). Participants were significantly less breathy (lower H1–A1 and H1–A3) in the attractive condition, supporting Hypothesis 1 and 2. In terms of vocal tract length, participants showed narrower formant dispersion in the attractive condition, signaling a larger body, rejecting Hypothesis 3. Their median f_o was also significantly lower when addressing an attractive face, thus

TABLE 3 | Summary of differences (attractive less unattractive facial stimuli) in each acoustic correlate by speakers' satisfaction in their own height (H being the most satisfied).

	Attractive *less* unattractive facial stimuli			
	H1–A1	H1–A3	Jitter	Shimmer
H	**−2.308**	**−2.627**	**27.447**	**−0.025**
M	−0.781	−1.773	18.349	−0.020
L	−0.787	−0.253	24.565	−0.016

Significant fixed effects are in bold.

rejecting Hypothesis 4. For creaky phonation, in the attractive face condition there was greater jitter but smaller shimmer, thus supporting Hypothesis 5 but not Hypothesis 6. For all voice quality-related measurements (i.e., analyses of breathiness and creakiness), there was a significant interaction between facial attractiveness rating and desired change in height.

Body Size Projection

As reported in Xu et al. (2013), male English listeners judged small-sounding acoustic cues to be more attractive, so even with cross-linguistic variation one would have expected Cantonese women to at least use some small-sounding cues in their production. Rather unexpectedly, in our data participants seemed to be always trying to project a large-sounding voice instead when speaking to an attractive face, unlike what the body size projection account would have predicted. This is reminiscent of the prevalent use of creaky voice by female American speakers, despite that creakiness is considered unattractive (Anderson et al., 2014). The case of creaky voice in American female speech shows that speakers do not necessarily use vocal strategies that listeners typically consider attractive – knowingly or otherwise. Another conceivable speculation is that speakers were taking into account social factors (classroom setting with friends nearby, interlocutor being a classmate), such that they deliberately avoided sounding too eager in front of an attractive potential mate. This speculation, needless to say, needs to be carefully verified.

In our initial analysis, we had the impression that speakers' height might affect their vocal strategies – this was confirmed in **Table 3**. For all voice quality-related acoustic cues, participants who were satisfied with their own height manifested a larger contrast between the attractive and the unattractive stimulus conditions. Our data thus seem to suggest that although female Cantonese speakers have the same set of vocal strategies for attractive vs. unattractive mates, it is those who are confident in their own height that are using them more actively.

Caveats

Participants in this study were well-educated young women who had been exposed to western culture since a very young age. They also spoke fluent English and Mandarin as second languages. This group of speakers thus represents only a subset of the local population. It is also noteworthy that when they took part in the production task, they had already been primed to think about attractiveness during the rating task – this could possibly have affected how they spoke. Finally, as is clear from **Supplementary Figure 2**, there is substantial individual variability in terms of vocal strategies. Therefore, this study may benefit from a larger sample than 19 speakers.

Suggestions for Future Research

Future studies should look at other groups of speakers in the community, such as older monolingual speakers. Another potentially interesting factor to investigate would be the effect of menstrual cycle on speech production. To the best of our knowledge, to date there is only preliminary data on how the menstrual cycle affects voice quality in Cantonese women (Li, 2016). Understanding how physiological factors interact with vocal attractiveness would shed new light on this issue. Finally, it would also be useful to verify the present findings with articulatory data, such as electroglottography (i.e., laryngograph).

CONCLUSION

This pilot study has found that young Cantonese women projected a large-sounding voice when speaking to an attractive male face. This seems to disagree with the widely held body size projection principle which states that an attractive female voice is small-sounding. We also found that women who were confident in their own height adjust their voice more actively depending on the attractiveness of their mates. Further investigation is needed to understand the relationship between the present findings and those observed in other languages.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by School of Humanities, the University of Hong Kong. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

EN conceived of the study supervised by AL. EN collected the data. AL analyzed the data and wrote the final manuscript. Both authors contributed to the article and approved the submitted version.

ACKNOWLEDGMENTS

An earlier version of this study appeared in Lee and Ng (2019).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2021.786507/full#supplementary-material

REFERENCES

Albert, G., Arnocky, S., Puts, D. A., and Hodges-Simeon, C. R. (2021). Can listeners assess men's self-reported health from their voice? *Evol. Hum. Behav.* 42, 91–103. doi: 10.3109/09638288.2013.793750

Anderson, R. C., Klofstad, C. A., Mayew, W. J., and Venkatachalam, M. (2014). Vocal fry may undermine the success of young women in the labor market. *PLoS One* 9:e97506. doi: 10.1371/journal.pone.0097506

Belin, P. (2021). "On voice averaging and attractiveness," in *Voice Attractiveness: Studies on Sexy, Likable, and Charismatic Speakers*, eds B. Weiss, J. Trouvain, M. Barkat-Defradas, and J. J. Ohala (Berlin: Springer), 139–149.

Berscheid, E., Dion, K., Walster, E., and Walster, G. W. (1971). Physical attractiveness and dating choice: a test of the Matching Hypothesis. *J. Exp. Soc. Psychol.* 7, 173–189. doi: 10.1016/0022-1031(71)90065-5

Chung, P.-P., and Leung, K. (1988). Effects of performance information and physical attractiveness on managerial decisions about promotion. *J. Soc. Psychol.* 128, 791–801. doi: 10.1080/00224545.1988.9924557

Coetzee, V., Greeff, J. M., Stephen, I. D., and Perrett, D. I. (2014). Cross-cultural agreement in facial attractiveness preferences: the role of ethnicity and gender. *PLoS One* 9:e99629. doi: 10.1371/journal.pone.0099629

Fitch, W. T. S. (1997). Vocal tract length and formant frequency dispersion correlate with body size in rhesus macaques. *J. Acoustical Soc. Am.* 102, 1213–1222. doi: 10.1121/1.421048

Gordon, M. K., and Ladefoged, P. N. (2001). Phonation types: a cross-linguistic overview. *J. Phonet.* 29, 383–406. doi: 10.1006/jpho.2001.0147

Gussenhoven, C. H. M. (2016). Foundations of intonational meaning: anatomical and physiological factors. *Top. Cogn. Sci.* 8, 425–434. doi: 10.1111/tops.12197

Hughes, S. M., and Miller, N. E. (2016). What sounds beautiful looks beautiful stereotype: the matching of attractiveness of voices and faces. *J. Soc. Pers. Relationships* 33, 984–996. doi: 10.1177/0265407515612445

Jäckle, S., Metz, T., Wenzelburger, G., and König, P. D. (2020). A catwalk to congress? appearance-based effects in the elections to the U.S. house of representatives 2016. *Am. Politics Res.* 48, 427–441.

Kreiman, J. E., Gerratt, B. R., and Antoñanzas Barroso, N. S. (2007). Measures of the glottal source spectrum. *J. Speech Lang. Hear Res.* 50, 595–610. doi: 10.1044/1092-4388(2007/042)

Kuznetsova, A., Brockhoff, P. B., and Christensen, R. H. B. (2017). lmerTest package: tests in linear mixed effects models. *J. Statist. Software* 82, 1–26.

Langlois, J. H., and Roggman, L. A. (1990). Attractive faces are only average. *Psychol. Sci.* 1, 115–121.

Lee, A., and Mok, P. K. P. (2021). "Lexical tone," in *The Cambridge Handbook of Phonetics*, eds R.-A. Knight and J. Setter (Cambridge: Cambridge University Press), 185–208.

Lee, A., and Ng, E. (2019). "Vocal attractiveness in cantonese: a production study," in *Proceedings of the 19th International Congress of Phonetic Sciences (ICPhS 2019)*. (Canberra, Act: Australasian Speech Science and Technology Association Inc).

Li, C. H. (2016). Voice onset time in Cantonese women across the menstrual cycle. *J. Acoustical Soc. Am.* 140:3112.

Louie, K. H. (2010). "Hong Kong on the move: creating global cultures," in *Hong Kong Culture: Word and Image*, ed. K. H. Louie (Hong Kong: Hong Kong University Press).

Morton, E. S. (1977). On the occurrence and significance of motivation-structural rules in some bird and mammal sounds. *Am. Nat.* 111, 855–869. doi: 10.1016/j.beproc.2009.04.008

Redi, L., and Shattuck-Hufnagel, S. R. (2001). Variation in the realization of glottalization in normal speakers. *J. Phonet.* 29, 407–429. doi: 10.1080/02699200802394856

Sarason, B. R., Sarason, I. G., Hacker, T. A., and Basham, R. B. (1985). Concomitants of social support?: social skills, physical attractiveness, and gender. *J. Pers. Soc. Psychol.* 49, 469–480. doi: 10.1037/0022-3514.49.2.469

Thornhill, R., and Gangestad, S. W. (1999). Facial attractiveness. *Trends Cogn. Sci.* 3, 452–460.

Watkins, L. M., and Johnston, L. (2000). Screening job applicants: the impact of physical attractiveness and application quality. *Int. J. Select. Assess.* 8, 76–84. doi: 10.1111/1468-2389.00135

Winkielman, P., Halberstadt, J., Fazendeiro, T., and Catty, S. (2006). Prototypes are attractive because they are easy on the mind. *Psychol. Sci.* 17, 799–806. doi: 10.1111/j.1467-9280.2006.01785.x

Xu, A., and Lee, A. (2018). Perception of vocal attractiveness by native Mandarin listeners. *Speech Prosody* 2018, 344–348.

Xu, A., Leung, S.-S., and Lee, A. (2017). Universal vs. language-specific aspects in human vocal attractiveness: an investigation towards Japanese native listeners' perceptual pattern. *Proc. Meet. Acoustics* 29, 1–8.

Xu, Y. (2013). "ProsodyPro: a tool for large-scale systematic prosody analysis," in *Proceedings of Tools and Resources for the Analysis of Speech Prosody (TRASP 2013)*. (Aix-en-Provence).

Xu, Y., Lee, A., Wu, W.-L., Liu, X., and Birkholz, P. (2013). Human vocal attractiveness as signaled by body size projection. *PLoS One* 8:e62397. doi: 10.1371/journal.pone.0062397

Yuasa, I. P. (2010). Creaky voice: a new feminine voice quality for young urban-oriented upwardly mobile American women? *Am. Speech* 85, 315–337. doi: 10.1215/00031283-2010-018

Zuckerman, M., and Driver, R. E. (1989). What sounds beautiful is good: the vocal attractiveness stereotype. *J. Nonverb. Behav.* 13, 67–82. doi: 10.1016/j.cub.2009.11.034

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Lee and Ng. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Psychology

ORIGINAL RESEARCH
published: 20 January 2022
doi: 10.3389/fpsyg.2021.822300

A Study on the Correlation Between Media Usage Frequency and Audiences' Risk Perception, Emotion and Behavior

*Peng-Peng Li and Fangqi Zhong**

College of Communication, National Chengchi University, Taipei City, Taiwan

Whether risk events can be effectively controlled and mitigated is largely influenced by people's perceptions of risk events and their behavioral cooperation. Therefore, this study used a web-based questionnaire (N = 306) to investigate the specific factors influencing people's risk perceptions and behaviors, and included a test for the difference in the effect of positive and negative emotions of the audiences. The results show that the overall model has good explanatory power (R^2 = 61%) for the behavioral variables, and (1) how people's use of different media (especially TV and online media) significantly influenced their positive and negative emotions; (2) how people's frequency of TV use significantly influenced their risk susceptibility and how online media use significantly influenced their risk severity (with some differences in people's perceptions of efficacy between different media); (3) how people's sense of efficacy for risky events is the strongest predictor of their risk control behavior; and (4) that there are different mediating effects of different emotions and risk severity and sense of efficacy between the frequency of media use and risk control behavior.

Keywords: emergent risk events, frequency of media use, emotion, risk perception, risk control behavior

OPEN ACCESS

Edited by:
Francesca D'Errico,
University of Bari Aldo Moro, Italy

Reviewed by:
Zira Hichy,
University of Catania, Italy
Haywantee Ramkissoon,
University of Derby, United Kingdom

***Correspondence:**
Fangqi Zhong
107463507@nccu.edu.tw

Specialty section:
This article was submitted to Personality and Social Psychology, a section of the journal Frontiers in Psychology

Received: *25 November 2021*
Accepted: *24 December 2021*
Published: *20 January 2022*

Citation:
Li P-P and Zhong F (2022) A Study on the Correlation Between Media Usage Frequency and Audiences' Risk Perception, Emotion and Behavior. Front. Psychol. 12:822300. doi: 10.3389/fpsyg.2021.822300

INTRODUCTION

In December 2019, a novel coronavirus (SARS-CoV-2) emerged and was officially named by the World Health Organization (WHO) as the novel coronavirus disease COVID-19 (Coronavirus Disease-2019) in February 2020 (Zhou et al., 2020). The World Health Organization characterized the new coronavirus outbreak as a public health emergency of international concern and declared the outbreak a "Global Pandemic" on March 11, 2020. In the face of this global catastrophe, effective risk communication and management, and the promotion of risk awareness and active prevention and control behaviors among the public to reduce the risk of disease transmission, have become important issues in risk communication.

Although governments and risk management units have attempted to prevent and control the development and spread of the pandemic through various measures, such as requiring masks in public places, social distancing, providing risk protection advice, and completely blocking travel for residents (Lundgren and McMakin, 2018), effective risk assessment and management programs had to take into account the public's perceptions and practical concerns, as risk assessment by technical experts is often based on rigorous theoretical constructs, experimental tests, and scientific evidence,

whereas public risk perceptions and assessments rely more on personal values, experiences, and subjective judgments.

Public risk perceptions are heavily influenced by media coverage (Singer and Endreny, 1993; Garfin et al., 2020), whether through traditional media or online social media (Muñiz, 2011; Yoo, 2019; Huynh, 2020), or interpersonal communication (Coleman, 1993; Rogers, 2000). When faced with the risk of a life-threatening disease, the public relies on the media to obtain accurate and up-to-date risk information in order to make informed decisions about health-protective behaviors. The "Media-System Dependency Theory" (Ball-Rokeach and DeFleur, 1976) suggests that, in times of risk crisis with high uncertainty, the public increases its reliance on the media (Casero-Ripollés, 2020; Huynh, 2020; Muñiz, 2020) and tends to obtain risk assessment and risk response advice from the media they perceive as trusted (Lachlan et al., 2016).

Previous studies have generally confirmed the influence of traditional media (Loges, 1994), including television, newspapers, and magazines, on people's risk perceptions, and in the modern era of widespread Internet access and rapid technological development, online social media have become an important channel for obtaining risk information (Karasneh et al., 2020; Malecki et al., 2020). However, it has also been suggested that, in this epidemic, people's trust in traditional media (especially television) has increased and they are more likely to obtain risk information from traditional media (Casero-Ripollés, 2020). At the same time, people's trust in and use of different media can significantly influence their emotions and risk perceptions (Coleman, 1993). The "risk as feeling hypothesis" suggests that people follow both cognitive (rational system) and emotional (experiential system) paths when assessing risk, and that emotions generally exert more influence on subsequent attitude formation and behavior (Loewenstein et al., 2001). For example, D'Errico and Paciello (2018, 2019) investigated the effect of people's emotions and cognitive processes on their online comments on specific issues; and in a subsequent study, they discussed the relevance of emotions or/and cognitive processes, and behaviors, and identified the effect of negative emotions based on moral or ethical dimensions, and of self-conscious emotions based on perceptions of social (collective) norms on individuals' pro-social behaviors (Paciello et al., 2021). In a study by Lei et al. (2014), it was suggested that the analysis of people's (online) emotions can help to understand the diversity of cognitive responses of different individuals to specific issues and can be used to develop better marketing or communication strategies. Focusing on risk communication and management, Ning et al. (2020) examined and confirmed the importance of people's risk perceptions, emotions, and knowledge and media messages on citizen protection behaviors during the COVID-19 outbreak through the knowledge-attitudes-practices (KAP) model. In other words, people's emotions and their cognitive processes about specific issues can be considered as effective and attractive communication signals across social and cultural contexts, as well as business and risk management domains. In view of this, the exploration of the frequency of people's media use under unexpected risk events and their emotional and behavioral reactions can identify promising media communication paths for effective risk communication.

In a study by Mejia et al. (2020), it was concluded that online social media and television representations of the risk of an epidemic are likely to trigger fear. Mertens et al. (2020) also concluded that fear is stronger among people who frequently use online social media to access risk information. However, although fear appeals have been shown in many studies to contribute to environmentally friendly behavior, excessive fear appeals may have negative effects (Witte, 1992; Witte et al., 2001; O'Neill and Hulme, 2009). In addition, Taha S. et al. (2014) found that during the H1N1 pandemic, anxiety was elevated when people's perceptions of uncertainty and uncontrollability of risk information content increased. In turn, anxiety motivates people to follow preventive behaviors recommended by risk management units, such as washing hands, wearing masks, and cleaning contact surfaces (Leung et al., 2003, 2004, 2005; Quinn et al., 2009). However, most of the previous studies on risk communication and people's perceptions and behaviors have focused on the effects of negative emotions, and the effects of positive emotions such as optimism and hope on risk-responsive behaviors have rarely been discussed. In addition to further examining the role of negative emotions in the epidemic, this study observes and discusses the influence and predictive power of positive emotions on people's risk perceptions and behaviors during the epidemic in order to improve the model and empirical study of emotions, risk perceptions, and risk transmission.

In summary, the two main factors that influence people's risk control behavior include emotions and risk perceptions, and the frequency of people's media use can influence these perceptions. Therefore, this study investigated the correlation between people's media usage frequency, emotions, and risk perceptions and behaviors using a web-based questionnaire, and then used a partial least squares approach to model the variance-based structural equation to assess the overall model's fitness and stability, providing a reference framework for effective risk communication based on the psychological aspects of people's emotions and risk perceptions.

LITERATURE REVIEW

Frequency of Media Use and Emotions

The Media-System Dependency Theory (MSD) (Ball-Rokeach and DeFleur, 1976) suggests that uncertain and ambiguous events or issues in society lead people to rely more on the media to resolve these ambiguous messages (Ball-Rokeach et al., 1984), and that when a major risk crisis occurs, the reliance on the media increases, and people are more likely to seek relevant information from the that than from interpersonal sources (Hirschburg et al., 1986). Similarly, Nikolaus Jackob theorized that that individuals will expose themselves to a trusted media environment (Jackob, 2010). In a study by Bangerter et al. (2012), it was confirmed that people's media dependence positively affects media trust, which in turn can influence people's behavioral intentions and the success of government-related policies. If people are able to obtain timely and accurate information from the medium they rely on, they will

trust the medium more and continue to rely on it. Conversely, if the information provided by the medium is uncertain and ambiguous, people will trust the medium less and turn to other sources to obtain relevant information.

Before discussing the relevance of media use to emotions, it is necessary to clarify the core distinction between emotions and affect. In general, emotional responses tend to be innate, occurring automatically and quickly (Cho et al., 2014). The difference between affect and emotion is that affect is "an assessment of the overall goodness or badness of an event or issue, focusing primarily on intuitive feelings" (Slovic et al., 2004), which is the broadest definition of the term and can include moods, feelings, and emotions at the core (Damasio, 1998); whereas emotion is a series of responses to specific stimuli, such as concepts, images, and objects, from the mind or different parts of the mind to the body, resulting in changes in perception and behavior within the body (Damasio, 1998; Kittipongvises and Mino, 2013; Tsung-Jen, 2017a). Compared to the more macroscopic concept of emotions, emotions have a higher substantive meaning and can represent specific reactions of individuals to different risk issues, situations, media reports or possible behavioral consequences, such as fear, and worry (Böhm, 2003).

In the literature on listeners' media use and their emotional responses, Ball-Rokeach and DeFleur (1976) concluded from the perspective of media dependence that listeners' use of different media can trigger and change their emotional responses to issues. Coleman (1993) and Garfin et al. (2020) confirmed that the type and frequency of media use affects people's psycho-emotional and physical reactions. Tracing their theoretical roots, the differential-impact hypothesis suggests that the content and function of different media (e.g., informational vs. entertainment media) exhibit different effects on the audience (Tyler and Cook, 1984; Glynn and Ostman, 1988; Coleman, 1993). As Tyler and Cook (1984) found, informational media (e.g., traditional news media) may influence individuals' rational analyses and judgments about the risks or responses to risks faced by others or society at large, while Snyder and Rouse (1995) showed that media with rich, diverse, and lively content (e.g., entertainment media) are more likely to stimulate individuals' emotional perceptions of what risks they will be exposed to. Holman et al. (2014), in a study examining media coverage and people's emotional reactions in the aftermath of the Boston Marathon bombing, noted that those with high media usage experienced higher levels of "acute stress disorder,"[1] which usually manifests as more negative emotions, such as fear, anxiety, and depression. Furthermore, Ramkissoon (2020, 2021a,b) emphasizes that family and friends or broader social network interactions increase people's positive emotions and help people improve their sense of efficacy by observing the effectiveness of collective behaviors, which in turn leads to proper risk recognition and response. Accordingly, this study proposes:

H1:The more people receive information about the epidemic through television (H1A), online media (H1B), and interpersonal (H1C), the stronger their negative emotional responses.

H2:The frequency of people's use of television (H2A), online media (H2B), and interpersonal (H2C) is significantly and positively correlated with their positive emotions.

Frequency of Media Use and Risk Perception

Risk perception involves people's subjective judgments about risks and benefits, which encompasses beliefs, attitudes, and other broader cultural values and social dispositions (UK Royal Society, 1992). In measuring people's subjective risk perceptions, the Health Belief Model (Janz and Becker, 1984) and the Extended Parallel Process Model (Witte, 1992) of health behavior theory can be divided into three main influences: perceived susceptibility, which is the likelihood of being affected by a risky hazard; perceived seriousness, which is the severity of the impact of a risky hazard; and self-efficacy, which is an individual's subjective behavioral judgment of his or her ability to perform or cope with a given task or situation (Bandura, 1986).

Harper et al. (2020) showed that people are more likely to engage in preventive behaviors when they perceive the risk to be serious and feel easily threatened. In a related study on efficacy, Tsung-Jen (2017b) mentioned that people assess not only their ability to act, but also the effectiveness of methods to solve problems, which is called "response efficacy." In a subsequent discussion of the social cognitive dimension, Bandura et al. (1999) and others also referred to "collective efficacy," the confidence of individuals knowing that collective action can solve common human problems.

The emergence of epidemics is not only caused by individuals, but also involves broader social and economic dimensions, and to some extent has connotations of collective action. Therefore, it may be more useful to consider collective efficacy along with self-efficacy and response efficacy to address high infectious risk diseases or other risk issues.

From the perspective of risk amplification by media, some scholars have suggested that the Internet can significantly enhance people's risk perceptions and behavioral responses (Kasperson et al., 1988; Skarlatidou et al., 2012), while Morton and Duck (2001) found in their study of health risk information dissemination on risk perceptions of listeners that, for those who use traditional media to obtain risk information, traditional media directly affects their own risk perceptions, while interpersonal communication indirectly affects their own risk perceptions. Ramkissoon (2021b) mentioned that interpersonal communication directly affects people's assessment of the severity of risk and their own sense of efficacy in coping with it. Some scholars have also argued that interpersonal communication enhances people's risk perceptions only when the mass media is not reporting, or is balanced in its coverage of a risk issue. It can be seen that the correlation between the degree of people's use of different media and their risk perceptions is still controversial, so:

[1]Acute stress disorder is defined as a reaction of fear, helplessness, or terror that reaches pathological levels after an individual experiences a traumatic stressful event (Department of Psychiatry, National Taiwan University School of Medicine Hospital, 2015).

RQ1: What is the difference between the frequency of people's use of TV, online media, and interpersonal and their effects on risk susceptibility, severity, and efficacy?

Emotion and Risk Perception

Empirical studies by Paul Slovic and others have shown that emotions are the main determinant of risk perception, and this pathway is considered to be the "emotional heuristic" or "risk-as-feeling" cognitive model that distinguishes it from rational risk assessment and analysis (Slovic, 1999; Finucane et al., 2000; Slovic et al., 2002; Böhm, 2003). At the same time, Lerner and Keltner (2001) argue that specific emotions can play different roles in risk perception. Also, specific emotions can have different degrees of influence on behavior (Nerb and Spada, 2001).

Experts in the field of environmental risk research have generally confirmed the positive association between negative emotions and people's risk perceptions (e.g., Leiserowitz, 2006; Smith and Leiserowitz, 2012, 2014), such as the higher the negative emotion index, the more likely it is that the risk (climate change) is perceived to cause serious impacts, leading people to adopt relevant coping behaviors and support-related policies. Lerner et al. (2003) found that fear increases people's perceptions of risk and harm and motivates them to act to protect themselves from potential threats, as well as confirming a positive relationship between fear, risk perceptions, and preventive behavior. In addition to fear, other negative emotions, such as anxiety and worry, have also been shown to be strongly associated with people's risk perception and control behaviors (Leung et al., 2003, 2004, 2005; Quinn et al., 2009). However, the Extended Parallel Process Model (EPPM) suggests that the negative emotion of fear results in two opposing cognitive responses, either by acting to control the risk and harm or by refusing to act, ignoring the risk, and reducing one's perception of fear (Witte, 1992; Witte et al., 2001). Leppin and Aro (2009) suggested that moderate levels of risk perception are beneficial for people to effectively combat epidemics and adopt preventive health behaviors, whereas high perceptions of risk of disease infection may prevent people from adopting effective preventive behaviors. It is nevertheless clear that negative emotions are mostly positively associated with people's perceptions of risk susceptibility and severity, and significantly associated with risk prevention and control behaviors, but few studies have investigated the association between negative emotions and perceptions of efficacy, and positive emotions and risk prevention and control behaviors, so:

H3: The more negative the people's emotions about the risk event, the higher their perception of risk susceptibility (H3A) and severity (H3B).

H4: Negative emotions are significantly and positively associated with risk control and control behaviors.

RQ2: What is the association between negative emotion and sense of efficacy?

Despite previous risk communication studies emphasizing the effect of negative emotions on risk perceptions and behaviors, the influence of positive emotions on risk perceptions and behaviors has only recently been emphasized, and the amount of research is still insufficient. Ojala's (2012) study found that, compared to negative hope based on denial, positive hope (constructive hope) is more conducive to positive behavioral intentions in adolescents. Returning to the context of this risk event, Aslam et al. (2020) used sentiment analysis to analyze the headlines of several prominent global media outlets reporting on risk and found that, while 52% of the headlines evoked negative emotions (e.g., fear, sadness, anger), 30% of the headlines still evoked positive emotions (e.g., hope and confidence). It is clear that the media does not intend to stimulate people's positive emotions, but positive emotions are also an important persuasion strategy. This study, however, did not explore the association between emotions, risk perceptions, and risk control behaviors.

Ramkissoon (2021a, cited Fredrickson's (1998)) argument for positive emotions in his study on how to maintain global well-being and public health, which emphasizes that positive emotions are more effective than negative emotions in enhancing people's psychological resilience and motivated behavioral responses to threats. However, given the lack of empirical research on the association between positive emotions and risk perceptions and behaviors, it is important to note that positive emotions are more likely to be associated with risk perceptions and behaviors, this study asks:

RQ3: Can positive emotions significantly affect the risk perception of the reader?

RQ4: Can positive emotions significantly influence the risk control behavior of listeners?

Risk Perception and Behavior

Risk perception is a key determinant of the public's willingness to engage in health protective behaviors. Studies of public health campaigns suggested that people's perceptions of potential risk may influence their risk behaviors (Brewer et al., 2004); Yıldırım and Güler (2020) also suggested that people's subjective understanding of risk in the context of new, unobservable and unpredictable risk crises (e.g., COVID-19) may influence their behavior. In another study by Bish and Michie (2010), the researchers classified the public's risk-protective behaviors in response to pandemic diseases into three categories, namely, preventive, avoidant, and management of disease. Preventive behaviors include washing hands, covering coughs, and sneezes with tissues, cleaning touching surfaces, wearing masks, and receiving vaccinations. Avoidant behaviors include avoiding crowded places, avoiding public transportation, and following epidemic control measures. Disease management includes taking antiviral medication, seeking professional help, and using the telephone or Internet to consult about the disease.

Slovic (1987) confirmed that people are more likely to avoid catastrophic risks that may lead to immediate and widespread death and ignore slow-onset risk threats, even though such risks may cause the same or greater mortality than that caused by sudden-onset disasters. In other words, people's perceptions of susceptibility and severity of risk increase their motivation to

protect themselves, which in turn motivates them to adopt risk-averse behaviors to avoid or reduce risk (Neuwirth et al., 2000). It has been suggested that the higher people's perceptions of their own or their friends' vulnerability to infection, the more likely they are to adopt risk-averse behaviors, such as avoiding public places or events. It has also been suggested that the higher the perceived risk of infection to oneself or to one's family and friends, the more likely one is to adopt risk-averse behaviors, such as avoiding public places or events and avoiding travel to countries or regions with high rates of infection (Lau et al., 2003, 2004; Blendon et al., 2004; Brug et al., 2004; Ramkissoon, 2020; Zheng et al., 2021), maintaining good personal hygiene (Tang and Wong, 2005; Wong and Tang, 2005), disinfecting homes (Lau et al., 2003; Wong and Tang, 2005; Rubin et al., 2009), wearing masks (Lau et al., 2004; Tang and Wong, 2005; Wong and Tang, 2005), strengthening the body's immunity through diet and exercise (Tang and Wong, 2005; Wong and Tang, 2005), and vaccination (Barr et al., 2008; Ramkissoon, 2020; Qiao et al., 2021) to prevent or reduce the risk of infection.

However, earlier studies on SARS risk perception and behavior did not find significant associations between risk susceptibility and hand washing behavior (Lau et al., 2007) and other preventive behaviors (Quah and Hin-Peng, 2004; Leung et al., 2005). In addition, people actively seeking drug prophylaxis, as well as counsel from experts, family or friends when dealing with infectious health risks, can reflect more diversity of perceived needs and behavioral motivations at the psychological level of the reader, Interpersonal communication can also influence the effectiveness of risk management to some extent (Ramkissoon, 2020, 2021a,b) and can be considered as an important observation and an area that needs to be urgently explored under COVID-19 risk events (2021b). This study defines this type of behavior as "disease management" and uses the research model to observe the representativeness and effectiveness of this type of behavior. In summary, this study proposes that:

> H5:People's risk susceptibility (H3A), severity (H3B), and efficacy (H3C) were significantly and positively associated with their risk prevention and control behaviors.

Combining the above research hypotheses and questions, the initial model of this study is shown in **Figure 1**.

MATERIALS AND METHODS

Survey Instrument and Sample Structure

Because the web-based questionnaire format can collect a large number of samples in a short period of time and has the advantages of low cost, visualization, and ease of subsequent research analysis (Fowler, 2002; Vaughn and Turner, 2016), it is gradually becoming one of the important tools for survey research at this stage. However, web-based questionnaires have disadvantages, such as under-representation, lack of sampling frame, and voluntary sampling, which become key issues to be addressed when choosing this method for data collection. Couper (2000), Witte and Allen (2000), and Dillman (2011) suggest that the data collection process should reflect the heterogeneity of the sample, the sample size should be increased and compared with the existing database, and the data collected should be adjusted by using statistical weighting, while avoiding excessive inferences when conducting research analysis. Inferences can effectively improve the reliability and validity of the questionnaire.

In addition, in terms of research tools, web-based questionnaire systems have become more and more functional in recent years and can largely meet the needs of researchers for different types (e.g., surveys, experiments) of research. At present, the more popular web-based questionnaire systems in the academic field are SurveyMonkey (United States), Questionnaire Star (Mainland), and SurveyCake (Taiwan). These three questionnaire systems all have the core functions of design, collection, and analysis, however, users can pay SurveyMonkey to specify the group to be tested, but its test samples are mainly people in Europe and the United States, and the subscription fee is high. Questionnaire Star is a simplified Chinese interface, its functions and models are rich with a low subscription fee. SurveyCake emerged in the Taiwan market, and is mainly used in academic research in Taiwan. The Chinese version of SurveyCake has been widely used to survey people's perceptions of issues and behavioral intentions in mainland China because of its relatively low subscription cost and mature development. Therefore, SurveyCake was used as the online survey tool in this study.

The questionnaire was firstly published on Weibo, WeChat friends circle, Xiaohongshu, and other online media with a large number of users, followed by contacting community management committees, college and university teachers and disseminating the questionnaire link, inviting people aged 18 and above in mainland China to fill in the questionnaire voluntarily. The system prevents each IP address from filling in the questionnaire more than once (through the duplicate IP deletion function of Questionnaire Star). The survey was conducted starting on December 20, 2020 and ended on January 6, 2021 with a total of 306 valid questionnaires collected.

Among the valid sample, 172 (56.2%) were male and 134 (43.8%) were female. The age distribution was mainly concentrated between 18 and 55 years old and relatively evenly distributed, ranging from 26 to 30 years old (17.7%), 20 to 25 years old (16.9%), 31 to 35 years old (15.6%), 36 to 40 years old (14.2%), 41to 45 years old (13.5%). In terms of respondents' marital status, the majority of respondents were married, with 135 (44.1%), followed by 93 (30.4%) who were cohabiting, 65 (21.2%) who were unmarried, 7 (2.3%) whose spouses had died, and 6 (2.0%) who were divorced or separated. The education degrees of the respondents were bachelor (103, 33.7%), master (79, 25.8%), high school (57, 18.6%), junior high school (31, 10.1%), doctoral (29, 9.5%), and others (7, 2.3%). In terms of sample structure, the gender, age, and marital status are very similar to the recent demographic data in mainland China, which represents the representativeness and inferred universality of the online sample. Only "education" differs from the census database. Lion et al. (2002) and Clayton and Myers (2015) pointed out that education is one of the internal factors that influence people's risk-responsive behavior, so this

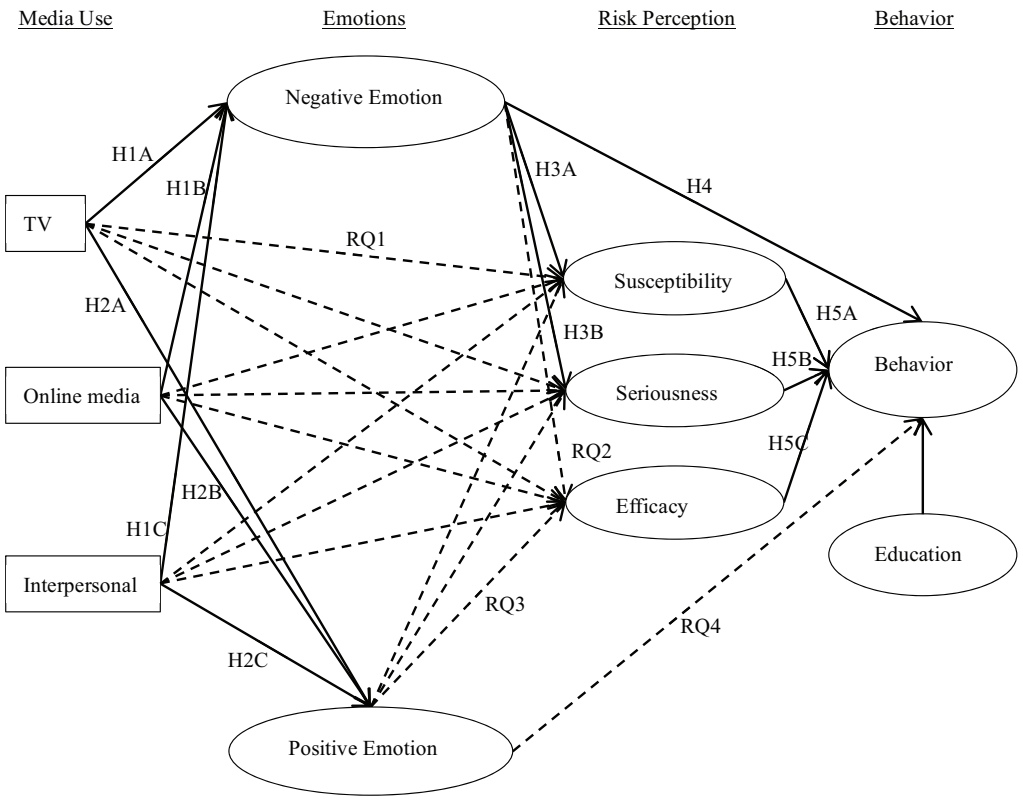


FIGURE 1 | Initial study model.

study set it as a control variable to observe its effect on the dependent variable.

Definition and Measurement of Variables

Frequency of Media Use Under Emergent Risk Events

The frequency of media use usually represents the exposure time and the degree of involvement and dependence of the reader on different media (Greenwald and Leavitt, 1984; Chaffee and Schleuder, 1986; de Vreese and Neijens, 2016). Mei-Ling and Yie-Jing's (2011) measurement of this variable specifically asks, "to what extent did you obtain information related to the outbreak during the New Coronary Pneumonia outbreak from the following media?" The options were (1) television, (2) online media, including WeChat, Weibo, and QQ, and (3) interpersonal, including family and friends, neighbors, and peers (colleagues), using a five-point scale [from 1 = never, 2 = rarely, 3 = sometimes (2–3 times a week), 4 = often (4–6 times a week), to 5 = almost every day].

Emotions

Emotions are a series of responses to specific stimuli, such as concepts, images, and objects, that move from the mind or different parts of the mind to the body, resulting in changes in perception and behavior within the body (Damasio, 1998; Kittipongvises and Mino, 2013; Tsung-Jen, 2017a). The scale used to measure the emotions of listeners in this study was based on Russell (1980, 2003), Russell and Barrett's (1999) "circumplex model," Scherer's (2005) "emotional dimensions scale," and Plutchik's (1984) "three-dimensional model of emotion structure," and was revised with reference to Briesemeister et al.'s (2011) emotion glossary, Mikels et al.'s (2005) and Stevenson et al.'s (2007) emotion measures, and the specificity of this study's topic. A five-point scale (1 = strongly disagree, 2 = disagree, 3 = unsure, 4 = agree, 5 = strongly agree) was used to ask: "How much do you agree with the following emotional adjectives in describing your emotional state during the New Coronary Pneumonia outbreak?" There were five adjectives representing positive emotions in the scale, namely optimistic, hopeful, confident, calm, and reassuring, and five adjectives representing negative emotions, namely fearful, depressed, scared, upset, and worried. Based on the overall model pretest results, adjectives with low factor loadings (<0.6 =) were removed (Hair et al., 1992), and finally, the observed variables of negative emotions were selected based on factor loadings > 0.6: fearful (0.82), worried (0.80), and anxious (0.78); the observed variables of positive emotion adjectives included: hopeful (0.88), confident (0.84).

Risk Perception

Referring to Rogers and Prentice-Dunn (1997) Health Belief Model (Janz and Becker, 1984) and the Extended Parallel Process Model (Witte, 1992) of protective motivation theory

and health behavior theory, three main potential variables were split to measure risk susceptibility perceptions, risk susceptibility perceptions, and efficacy (Frewer et al., 2002).

Perceptions of risk susceptibility (i.e., likelihood of being at risk of harm): (1) To what extent do you think you are susceptible to COVID-19, (2) To what extent do you think you can avoid getting COVID-19, and (3) To what extent do you think you are a potential factor in spreading COVID-19 (five-point scale: 1 = not at all, 2 = not, 3 = not sure/don't know, 4 = probably, 5 = to a large extent).

Sense of efficacy (includes confidence in individual and collective ability to cope with risk and the effectiveness of risk response programs): (1) Do you think you can do simple things to reduce the threat of NCCV infection; (2) Do you think government measures to prevent NCCV infection (e.g., wearing masks, washing hands, keeping social distance, etc.) can reduce the spread of NCCV and the chance of infection; (3) Do you think that the measures provided by the government to prevent the spread of neoplasmosis (such as wearing a mask, washing hands regularly, keeping social distance, etc.) can reduce the chance of transmission and infection of neoplasmosis (5-point scale: 1 = not at all, 2 = not at all, 3 = not sure/don't know, 4 = yes, 5 = to a large extent).

Risk Control Behavior

Referring to Bish and Michie's (2010) analysis, this study classified behavioral response intentions into three categories: risk prevention, risk avoidance, and disease management, and asked participants on a five-point scale (1 = definitely not, 2 = not, 3 = not sure, 4 = probably, 5 = definitely) how likely they would be to adopt the following behaviors during a new coronary pneumonia outbreak, including:

> Risk prevention behaviors: (1) wearing a mask outside; (2) washing hands regularly; (3) covering up with a tissue when coughing or sneezing; (4) cleaning the surfaces of touching objects.
>
> Risk avoidance behaviors: (1) avoid going out; (2) if you must go out, avoid taking public transportation; (3) if you must go out, avoid crowded places; (4) avoid buying food for delivery.
>
> Disease management behaviors: (1) taking antiviral medications to prevent viral infection; (2) using the telephone or Internet to consult professionals about symptoms of new coronary pneumonia or prevention advice; (3) proactively discussing disease related information with friends, relatives or others through the Internet, social media or face-to-face.

In order to simplify the research model as much as possible, these 11 behavioral items were first analyzed by principal component factor analysis, and then were extracted by Bartlett's test of sphericity using the eigenvalue greater than one rule. The total variance was 77.13% explained by the three factors (i.e., the observed constructs of risk prevention, avoidance, and disease management behavior intention). Also, the factor loadings of the 11 items of the scale were all greater than 0.7, which has high convergent validity (Hair et al., 1992). Therefore, in the formal analysis of this study, the scores of the three types of behavioral items were summed, with higher scores indicating higher willingness to respond to the behavior in question.

Non-response Bias

Referring to Armstrong and Overton's (1977) and Shiau and Luo's (2012) test procedures for non-response bias, this study examined the issue of non-response bias by comparing the gender and age of early (N = 187) and late (N = 119) participants. The chi-square test for early and late participants showed no significant difference in gender or age ($p > 0.05$). Therefore, the possibility of non-response bias was ruled out.

Common Method Bias

In this study, the questionnaire was designed using the "Questionnaire Star" software and the data were distributed and collected through an online questionnaire. In order to ensure the validity of the follow-up analysis, on the one hand, the questionnaire was designed in accordance with previous research theories and implementation methods. The main purpose of the study was described to the participants, ethics and precautions were informed, the questions were filled in anonymously, and the questions of each variable were randomly ordered to minimize the influence of CMB on the results of the study. On the other hand, we used a *post hoc* test (Lindell and Whitney, 2001) in the later screening of the valid sample. Following the procedure of Hultman et al. (2009), we isolated the observed variables with a minor positive correlation (0.015) and then calculated the CMB-adjusted correlation between all latent variables and found that the difference between the original and adjusted correlation coefficients was very small ($r \leq 0.015$) and the significant correlation of the variables did not change. At the same time, the chi-square test of fit showed no significant results ($X^2 = 17$, $p > 0.05$), so the CMB had no significant effect on the results and inferences of this study.

Partial Lease Square-Structural Equation Modeling Method

Structural equation modeling (SEM) allows us to assess the correlation and theoretical connection of one or more independent variables with one or more dependent variables in terms of discrete or continuous variables (do Nascimento and da Silva Macedo, 2016). Its advantage lies mainly in the ability to effectively analyze the causal relationships between multiple variables (Williams et al., 2009; Durdyev et al., 2018). There are two standard approaches to SEM, namely Covariance Base (CB-SEM), based on covariance, often using LISREL, EQS, and AMOS software, and partial lease square (PLS), which analyzes the principal component structure of variables. In the past decades, CB-SEM has been considered the main way to analyze the complex relationship between observed and latent variables, while PLS-SEM has become more popular in recent years in social science research. Both have unique advantages in terms of analytical sample structure, objectivity, and model fitness, but the use of PLS-SEM is more appropriate when the research goal is to develop a theory with a small sample size and a more complex

model (Hair et al., 2019; Khan et al., 2019; Shiau et al., 2019; Wen-Long, 2020). On the contrary, if the study aims to verify or confirm theoretical connections, CB-SEM is the most appropriate method (do Nascimento and da Silva Macedo, 2016; Peng et al., 2019).

PLS-SEM is more suitable for this study than CB-SEM for the following reasons: (1) this study is an exploratory study aimed at testing and developing existing theories and research results; (2) the analytical results are used to predict variable correlations; (3) there are more variables and the overall model is more complex; (4) the effective sample size collected is smaller; and (5) the sample distribution lacks normality (Gefen et al., 2011; Hair et al., 2019; Khan et al., 2019; Shiau et al., 2019; Wen-Long, 2020). Therefore, in this study, we ran SmartPLS 3 to test the reliability and validity of this structural model by plotting the model and validating the factor analysis. **Table 1** shows that all the factors in this study have a good convergent validity with factor loadings > 0.6 (Hair et al., 1992; Wixom and Watson, 2001) and $p < 0.01$ (Bock et al., 2005). Each construct has an AVE indicator > 0.05 (Chin et al., 2003). This means that the factors in this study met the stability criteria.

Next, this study examined the correlation and extracted average variance (AVE) of the potential variables to confirm the discriminant validity of the observed items (Gefen et al., 2000). The results showed the discriminant validity of the potential variables for subsequent analysis and inference (see **Table 2**).

STRUCTURAL MODELING ANALYSIS

After testing the reliability and validity of the initial research model, this study proceeded to perform path analysis and correlation testing of the overall model using PLS. The path coefficients and significance of each research hypothesis in the structural model were verified by a bootstrap resampling method (1,000 resamples), and the results are shown in **Figure 2**.

From the results of the analysis, the model explained 34% of the variance in negative emotions, 29% of the variance in positive emotions, 31% of the variance in susceptibility, 40% of the variance in severity, 55% of the variance in efficacy, and 61% of the variance in behavior.

First, in terms of the correlation between people's media usage frequency and their emotions: (1) people's frequency of online media usage was significantly positively correlated with their negative emotions ($\beta = 0.62$, $p < 0.001$), i.e., the more people tended to obtain epidemic-related information from online media, the stronger their negative emotions were (H1B holds); (2) its significance exceeded the significant positive correlation between TV usage frequency and their negative emotions ($\beta = 0.37$, $p < 0.05$), i.e., the more people tended to obtain epidemic-related information from TV media, the stronger their negative emotions (H1A holds); (3) the significant positive correlation between frequency of interpersonal network use and negative emotions ($\beta = 0.24$, $p < 0.05$), i.e., the more people tend to obtain epidemic-related information from interpersonal media, the stronger their negative emotions (H1C holds). In terms of the correlation between people's frequency of use of different media and their positive emotions; and (4) the more often people obtained risk information through television ($\beta = 0.46$, $p < 0.001$) and online media ($\beta = 0.15$, $p < 0.05$), the stronger their positive emotions were (H2A, H2B), but the frequency of interpersonal use was not significantly correlated with positive emotions (H2C did not hold).

Secondly, from the statistical results of people's frequency of use of different media and their risk perceptions, although people's frequency of use of online media was not significantly correlated with their risk susceptibility, it was significantly correlated with people's risk severity and sense of efficacy, and its significance exceeded the effects of television and interpersonal on people's risk susceptibility and sense of efficacy (RQ1), specifically: (1) the higher the frequency of people's use of television, the higher their risk susceptibility ($\beta = 0.28$, $p < 0.05$), severity ($\beta = 0.23$, $p < 0.05$), and efficacy ($\beta = 0.40$, $p < 0.01$); (2) the more frequently people used online media, the higher their risk severity ($\beta = 0.71$, $p < 0.001$) and lower their efficacy ($\beta = -0.53$, $p < 0.001$); and (3) the more frequently people used interpersonal higher frequency, the higher their risk severity ($\beta = 0.19$, $p < 0.05$) and sense of efficacy ($\beta = 0.18$, $p < 0.05$).

Again, in terms of the association between people's emotions and risk perceptions and risk control behaviors, the statistical results showed that (1) people's negative emotional responses to the risk event were not significantly associated with their risk susceptibility ($\beta = 0.09$, $p > 0.05$) and severity ($\beta = 0.07$, $p > 0.05$) (H3 does not hold), nor with their sense of efficacy ($\beta = 0.10$, $p > 0.05$) [RQ2 (RQ2)]; (2) people's positive emotional response to the risk event was significantly and positively correlated with their sense of efficacy ($\beta = 0.33$, $p < 0.001$), i.e., the stronger people's positive emotions were, the higher their confidence in themselves, their government, and the group to fight the risk, but positive emotions were not significantly correlated with their risk susceptibility and severity (RQ3); (3) people's negative emotions were weakly and significantly positively correlated with their risk control behavior ($\beta = 0.19$, $p < 0.05$); and (4) people's negative emotions were weakly and positively correlated with their risk control behavior (RQ4) ($\beta = 0.19$, $p < 0.05$), i.e., moderate negative emotions help to stimulate positive risk prevention behavior (H4 holds). There was no direct significant relationship between people's positive emotions and their risk control behavior ($\beta = 0.04$, $p > 0.05$) (RQ4).

Finally, in terms of the association between risk perception and risk control behavior, the statistical results confirmed that people's risk susceptibility ($\beta = 0.17$, $p < 0.05$), severity ($\beta = 0.14$, $p < 0.05$), and efficacy ($\beta = 0.43$, $p < 0.05$) were significantly and positively associated with their risk control behavior, i.e., the more people perceived the risk as serious, their vulnerability to infection and injury, and their confidence in risk control, the more positive their risk prevention behavior was.

In addition, this study also examined the indirect pathway of media frequency on risk-causal behavior, i.e., whether emotions and risk perceptions mediated the effect of media frequency on risk-causal behavior. The results of the analysis, controlling for educational attainment variables, showed that

TABLE 1 | Scale properties of structure model.

Construct	Item	Mean	Standard deviation	Error loading	Standardized item loading	T-statistic
TV	TV	3.76	0.74	0.37	0.88	29.17***
Online media	Online media	3.91	0.69	0.29	0.86	32.33***
Interpersonal	Interpersonal	3.85	0.67	0.35	0.87	37.64***
Positive emotion	Promising	3.55	0.84	0.32	0.88	42.91***
	Confident	3.64	0.77	0.30	0.84	37.26***
Negative emotion	Fearful	3.99	0.61	0.29	0.82	25.77***
	Worried	3.97	0.66	0.26	0.80	39.97***
	Anxious	3.88	0.73	0.22	0.78	43.10***
Susceptibility	Sus1	3.47	0.80	0.24	0.78	39.95***
	Sus2	3.61	0.78	0.19	0.80	47.83***
	Sus3	3.42	0.81	0.27	0.88	53.21***
Seriousness	Ser1	3.88	0.68	0.24	0.83	48.75***
	Ser2	3.86	0.67	0.18	0.88	21.45***
	Ser3	3.75	0.73	0.19	0.86	51.00***
Efficacy	Se	3.39	0.81	0.34	0.81	44.17***
	Re	3.44	0.81	0.33	0.79	29.93***
	Ce	3.78	0.76	0.27	0.80	42.24***
Behavior	preventive	3.62	0.79	0.30	0.89	59.25***
	avoidant	3.87	0.70	0.24	0.87	45.69***
	management	3.73	0.79	0.33	0.85	36.27***

*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

TABLE 2 | Scale properties and correlations.

Construct	Number of items	Reliability	AVE	Factor correlations									
				TV	OM	IP	PE	NE	Sus	Ser	Ef	Be	Ed
TV	1	0.93	0.74	1.0									
OM	1	0.95	0.72	−0.01	1.0								
IP	1	0.90	0.77	0.05	0.12	1.0							
PE	2	0.88	0.70	0.62	0.28	0.14	1.0						
NE	3	0.90	0.78	0.33	0.44	0.37	0.13	1.0					
Sus	3	0.94	0.81	0.53	0.18	0.16	0.09	0.08	1.0				
Ser	3	0.92	0.83	0.47	0.50	0.28	0.04	0.14	0.16	1.0			
Ef	3	0.91	0.80	0.66	0.73	0.22	0.31	0.17	−0.18	−0.37	1.0		
Be	3	0.95	0.75	0.22	0.25	0.17	0.28	0.27	0.36	0.29	0.34	1.0	
Ed	6	0.93	0.86	0.70	0.69	0.66	0.55	0.25	0.44	0.36	0.29	0.61	1.0

OM, Online medi; IP, Interpersonal; PE, Positive Emotion; NE, Negative Emotion; Sus, Susceptibility; Ser, Seriousness; Ef, Efficacy; Be, Behavior; Ed, Education.

(1) negative emotions mediated the positive effect of TV (95% CI: 0.42–0.51, $P < 0.001$) and online media (95% CI:0.58–0.73, $P < 0.001$) usage frequency on behavior; (2) positive emotions mediated the positive effect of TV usage frequency on behavior (95% CI: 0.19–0.32, $P < 0.01$); (3) positive emotions mediated the positive effect of TV usage frequency on behavior. $P < 0.01$); (4) sense of efficacy mediated the positive effect of TV usage frequency on behavior (95% CI: 0.29–0.36, $P < 0.01$); (5) risk severity mediated the positive effect of interpersonal usage frequency on risk prevention behavior (95% CI: 0.31–0.45, $P < 0.05$); and (6) online media mediated the positive effect on behavior through both risk severity and sense of efficacy was positively influenced (95% CI: 0.17–0.31, $P < 0.01$).

DISCUSSION

This study investigated the factors that influence people to adopt positive risk prevention behaviors in the context of sudden risk events (COVID-19), and whether the frequency of using different media is significantly related to their emotions (positive and negative), risk perceptions (susceptibility, seriousness, and efficacy), as well as the mediating effect the paths between media frequency and risk-responsive behaviors may have.

Our findings indicate, first, that the degree (frequency) of people's use of different media in this COVID-19 risk event significantly affected their negative emotions of fear, worry, and anxiety, and the influence of online media on their negative emotions exceeded that of interpersonal and television, while

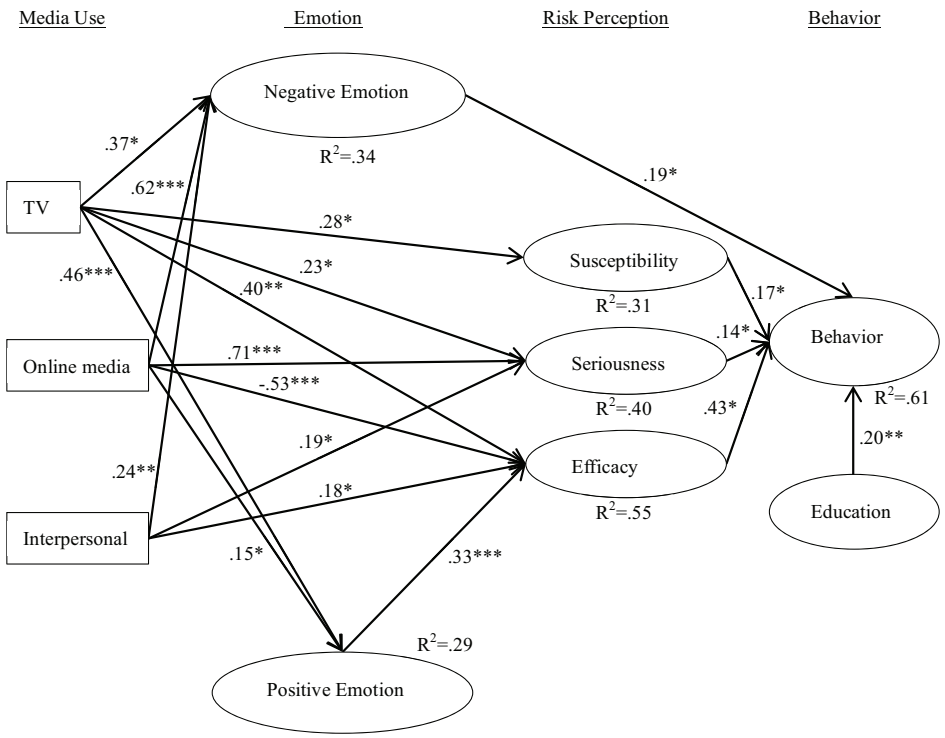


FIGURE 2 | Results of the structure model. $N = 306$, $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

positive emotions were mainly influenced by people's use of television. In contrast, online media showed a weak correlation with positive emotions. Possible explanations include:

(1) When a sudden risk event occurs, scientists and traditional media (e.g., television) have difficulty in reporting accurate risk information in a timely manner, because there is not yet sufficient scientific data to confirm the appearance of risk, its destructive power, and effective preventive measures (Henrion and Fischhoff, 1986), and people may become anxious and worried because of the uncertainty and ambiguity of risk information. When scientists and media have conclusive evidence of risk hazards and countermeasures, they still cannot change their emotions or even doubt the accuracy and credibility of the information (Nyhan and Reifler, 2010), and so turn to online media and people to obtain relevant information. However, because there are many unverified messages on the Internet, people's negative emotions intensify and spread.

(2) People may be influenced by their education level, risk knowledge, risk experience, and values to trust the risk information and response advice provided by the official media on TV, and then rationally analyze risks and make rational behavioral decisions. This finding is in line with Tyler and Cook (1984) analysis of the association between information media (traditional media) and people's rational analysis and assessment of risk, which suggests that traditional media are still effective in guiding people to respond rationally to risk. The trust of the population in traditional media (especially TV) has increased and they are more likely to obtain risk information from traditional media. In addition, due to the aggregation of information preferences of different people by big data, people who perceive things more rationally or objectively on a daily basis are more likely to see information content that is consistent with their perceived preferences when risks occur, such as scientific analysis of risks and how to prevent infection. However, these people are still a minority in society as a whole, so online media is less influential than television in terms of positive sentiment. In terms of the association between interpersonal communication and emotion, this study did not reach the same conclusion as Ramkissoon (2020, 2021a,b)) that social network interaction (interpersonal communication) would enhance people's positive emotion.

Second, people's perceptions of risk susceptibility are mainly influenced by their frequency of using TV, while their perceptions of risk seriousness are mainly influenced by their frequency of using online media, followed by TV and interpersonal. Their sense of efficacy varies depending on their frequency of using different media. This conclusion supports Ramkissoon's (2021b) comments on the importance of interpersonal communication, which significantly and positively affects people's perceptions of risk severity and efficacy. People's frequency of using TV and interpersonal enhances their sense of efficacy, while the higher the frequency of using online media, the lower their

sense of efficacy. This finding suggests that people's frequency of media use is indeed significantly related to their risk perceptions (Morton and Duck, 2001), but that the use and trust of different media have different effects on their risk susceptibility, seriousness, and efficacy.

Our third finding confirmed that people's perceived efficacy of the COVID-19 risk event was the strongest predictor of their risk-response behavior, followed by negative emotion, risk susceptibility, and seriousness in that order. This finding is consistent with the findings of numerous studies focusing on the fact that people are more active in risk prevention behaviors when they perceive the risk to be serious and feel easily threatened (Harper et al., 2020), and when they perceive that individual (Tsung-Jen, 2017a) and collective action (Bandura et al., 1999; Ramkissoon, 2021b) can effectively counteract risk harm (Ramkissoon and Smith, 2014; Ramkissoon, 2020).

As for mediating effects, our fourth finding shows that different emotions and risk seriousness and efficacy have different mediating effects between the frequency of media use and risk control behavior. Emotions and risk seriousness and efficacy are important mediating variables in people's processing of risk information. As shown in the two-path model proposed by Stúrmer and Simon (2004) and the findings of Tsung-Jen (2017a), people will act accordingly because of their negative emotions, due to the perceived relevance of risk to them, and will also enhance their sense of efficacy and behavioral intentions with the risk coping information obtained from the media.

The model explained 34% of the variance in negative emotions; 28% of the variance in positive emotions; 31% of the variance in susceptibility; 40% of the variance in severity; 55% of the variance in efficacy; and 61% of the variance in risk prevention and control behaviors. According to Cohen (1988) and Falk and Miller (1992), R Square should be equal to or greater than 0.10, which means that the variance explanation of the endogenous structure is sufficient, and the larger the value, the stronger the explanatory power of the antecedent variable for that variable. It can be seen that the structural model in this study has strong explanatory power for people's risk prevention behavior. Therefore, this study advises that the media should pay more attention to the presentation of risk information in future risk communication, fully consider the possible effects of risk information on people's emotions and risk perceptions, and avoid using strong negative emotional appeals (e.g., panic) and language that can be easily misunderstood (e.g., a word may have multiple meanings, some words may be used in a wide variety of ways, some words may change their meaning when combined with other words, or the meaning of a word may change), or ambiguous information (e.g., a word may mean different things to different people, such as "significant," Tucker and Ferson, 2008), to give people confidence in their ability to cope with risk and thus to adopt positive and effective risk prevention behaviors.

As Lundgren and McMakin (2018) suggest, demonstrating the ability of government and related entities to protect the public from risk or to provide sufficient empathy to the public in the dissemination of risk information and processes is the most critical element to enhance organizational trust and effective risk management. The significant impact of traditional media on the public's ability to correctly perceive and analyze risks and to make appropriate behavioral decisions should not be overlooked. On the one hand, considering the changing media habits of the public, government or relevant risk management units should make good use of online media, such as disseminating risk-related information in a timely manner, understanding the psychological state of the public and their general concerns and communicating with them effectively, and improving the presentation of risk information to the general public to facilitate their clear understanding and perception of risks.

On the other hand, due to the high accessibility and low threshold of online media, people are easily influenced by "news" that is mixed with the real and the fake, and so may spread the unchecked information to their personal social networks and larger media platforms, resulting in an increase of negative emotions, such as collective fear, worry, and anxiety (Taha S. A. et al., 2014; Jones et al., 2017). They may overestimate certain threat factors that are actually less risky or underestimate certain factors that are actually more risky, resulting in one-sided or erroneous risk perceptions (Slovic et al., 1980), and even disordered or overly aggressive risk responses and violent emotional outbursts, such as hearing rumors that spraying alcohol on masks can improve their protective ability. This is why government has to take a more active role in communication and management. This means that in the process of risk communication and management, government should improve the transparency of information and the speed of updating information, prevent people from speculating and spreading vague or uninformed risk information as much as possible, establish fact-checking platforms and mechanisms to stop the vicious spread of fake news in online media and among people, and strengthen the cultivation of media literacy among people in the field of education, so as to effectively manage risks and promote a rational society.

Previous health risk communication studies have mostly focused on protective motivation and health belief models, exploring the effects of perceived severity, perceived susceptibility, perceived benefits, and perceived barriers on risk prevention and treatment behaviors of listeners. However, COVID-19 is a global risk event that involves both environmental and health risks, which are not only related to health practices and health at the individual level, but also to the collective interests of society. This requires a more diversified exploration of the psychological factors that influence the behavior of listeners, as discussed in Loewenstein et al.'s (2001) risk-as-feelings hypothesis, which suggests that people will follow both cognitive and emotional paths to form specific attitudes and make behavioral decisions, which affirms the influence of emotions on attitudes and behaviors. However, subsequent research on the influence of emotions on risk-responsive behavior has focused on negative emotions (especially fear) and has failed to identify in detail the effects of which negative emotions and positive emotions influence behavior.

Among the many studies on negative emotions and behaviors, the Extended Parallel Process model (EPPM) suggests that listeners' processing of risk information can be divided into threat

assessment and effectiveness assessment (Witte, 1992; Witte et al., 2001). The threat assessment includes the listener's assessment of the severity and susceptibility of the risk, or the relevance of the risk, while the efficacy assessment includes the assessment of one's own and the group's ability to cope with the risk, and the effectiveness of risk advice. Combined with the Health Belief Model and the EPPM, this study classifies risk perception into three important dimensions: susceptibility, severity, and efficacy, and analyzes the differential effects of different (positive/negative) emotions and risk perception dimensions on the frequency of media use and the risk prevention and control behavior paths of the listeners. The results of this study are not only applicable to the field of health risk communication, but also have an important reference value in the field of environmental risk communication.

However, the types of media in modern society are very diverse, and it is difficult to clearly distinguish which media are traditional media or online media. It is suggested that, in the future, we can increase the number of observed variables of online media (e.g., official web pages, online search engines, outdoor online media, short video platforms, etc.) and dialogue with the results of this study to further enhance the explanatory validity and generalizability of this study's model. At the same time, in addition to the core variables discussed in this study, the factors influencing individuals' risk-responsive behavior should not neglect the guiding or moderating effects of individual values, risk knowledge, risk experience, social norms, and other socio-demographic variables (e.g., gender, age, residence, etc.) on their behavior (see Mary and Wildavsky, 1982; Morgan et al., 1985; UK Royal Society, 1992; Peters and Slovic, 2000; Edelstein, 2002; Slovic, 2016).

As suggested by Ramkissoon (2020, 2021a,b), place attachment and habitual behavior are also important factors influencing people's risk perceptions and behaviors, and have an important impact on people's well-being and health. Although this study has controlled for the influence of educational attainment on behavior, given the parsimonious nature of the overall research model, other factors influencing behavior have not been discussed comprehensively, and it is suggested that future research could add or change the control variables and add moderating variables to observe and test whether the explanatory power of this study's model and findings have changed and what differences exist.

In addition, considering the urgent need to maintain people's health and enhance human well-being in the context of modern risk society and globalization (2021b), future research may focus on the ecological environment and people's psychological recovery after the epidemic, and how to enhance people's happiness and well-being, such as building (Ramkissoon, 2021a) and promoting (Ramkissoon, 2020) emotional connections between individuals and activity sites, and the establishment and promotion of ecotourism to repair negative emotional stress (Majeed and Ramkissoon, 2020), and to compare the psychological and behavioral effects of different social institutions and cultural contexts on people's risk coping.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

P-PL: conceptualization, methodology, data curation, and writing-original draft preparation. FZ: supervision and writing-reviewing and proofing. Both authors contributed to the article and approved the submitted version.

REFERENCES

Armstrong, J. S., and Overton, T. S. (1977). Estimating nonresponse bias in mail surveys. *J. Mark. Res.* 14, 396–402. doi: 10.1177/002224377701400320

Aslam, F., Awan, T. M., Syed, J. H., Kashif, A., and Parveen, M. (2020). Sentiments and emotions evoked by news headlines of coronavirus disease (COVID-19) outbreak. *Humanit. Soc. Sci. Commun.* 7:23. doi: 10.1057/s41599-020-0523-3

Ball-Rokeach, S. J., and DeFleur, M. L. (1976). A dependency model of mass-media effects. *Commun. Res.* 3, 3–21. doi: 10.1177/009365027600300101

Ball-Rokeach, S. J., Rokeach, M., and Grube, J. W. (1984). *The Great American Values Test: Influencing Behavior and Belief Through Television*. New York, NY: Free Press.

Bandura, A. (1986). *Social Foundations of Thought and Action: A Social Cognitive Theory*, Vol. 1986. Englewood Cliffs, NJ: Prentice-Hall.

Bandura, A., Freeman, W., and Lightsey, R. (1999). *Self-Efficacy: The Exercise of Control*. New York, NY: Freeman.

Bangerter, A., Krings, F., Mouton, A., Gilles, I., Green, E. G., and Clémence, A. (2012). Longitudinal investigation of public trust in institutions relative to the 2009 H1N1 pandemic in Switzerland. *PLoS One* 7:e49806. doi: 10.1371/journal.pone.0049806

Barr, M., Raphael, B., Taylor, M., Stevens, G., Jorm, L., Giffin, M., et al. (2008). Pandemic influenza in Australia: using telephone surveys to measure perceptions of threat and willingness to comply. *BMC Infect. Dis.* 8:117. doi: 10.1186/1471-2334-8-117

Bish, A., and Michie, S. (2010). Demographic and attitudinal determinants of protective behaviours during a pandemic: a review. *Br. J. Health Psychol.* 15, 797–824. doi: 10.1348/135910710X485826

Blendon, R. J., Benson, J. M., DesRoches, C. M., Raleigh, E., and Taylor-Clark, K. (2004). The public's response to severe acute respiratory syndrome in Toronto and the United States. *Clin. Infect. Dis.* 38, 925–931. doi: 10.1086/382355

Bock, G.-W., Zmud, R. W., Kim, Y.-G., and Lee, J.-N. (2005). Behavioral intention formation in knowledge sharing: examining the roles of extrinsic motivators, social-psychological forces, and organizational climate. *MIS Q.* 29, 87–111. doi: 10.2307/25148669

Böhm, G. (2003). Emotional reactions to environmental risks: consequentialist versus ethical evaluation. *J. Environ. Psychol.* 23, 199–212. doi: 10.1016/S0272-4944(02)00114-7

Brewer, N. T., Weinstein, N. D., Cuite, C. L., and Herrington, J. E. (2004). Risk perceptions and their relation to risk behavior. *Ann. Behav. Med.* 27, 125–130. doi: 10.1207/s15324796abm2702_7

Briesemeister, B. B., Kuchinke, L., and Jacobs, A. M. (2011). Discrete emotion norms for nouns: Berlin affective word list (DENN–BAWL). *Behav. Res. Methods* 43, 441–448. doi: 10.3758/s13428-011-0059-y

Brug, J., Aro, A. R., Oenema, A., De Zwart, O., Richardus, J. H., and Bishop, G. D. (2004). SARS risk perception, knowledge, precautions, and information

sources, the Netherlands. *Emerg. Infect. Dis.* 10, 1486–1489. doi: 10.3201/eid1008.040283

Casero-Ripollés, A. (2020). Impact of Covid-19 on the media system. Communicative and democratic consequences of news consumption during the outbreak. *Prof. Inform.* 29:e290223. doi: 10.3145/epi.2020.mar.23

Chaffee, S. H., and Schleuder, J. (1986). Measurement and effects of attention to media news. *Hum. Commun. Res.* 13, 76–107. doi: 10.1111/j.1468-2958.1986.tb00096.x

Chin, W. W., Marcolin, B. L., and Newsted, P. R. (2003). A partial least squares latent variable modeling approach for measuring interaction effects: results from a Monte Carlo simulation study and an electronic-mail emotion/adoption study. *Inf. Syst. Res.* 14, 189–217. doi: 10.1287/isre.14.2.189.16018

Cho, H., Reimer, T., and McComas, K. A. (2014). *The Sage Handbook of Risk Communication*. Thousand Oaks, CA: SAGE Publications.

Clayton, S., and Myers, G. (2015). *Conservation Psychology: Understanding and Promoting Human Care for Nature*. Hoboken, NJ: Wiley-Blackwell.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*, 2nd Edn. New York, NY: Routledge. doi: 10.4324/9780203771587

Coleman, C. L. (1993). The influence of mass media and interpersonal communication on societal and personal risk judgments. *Commun. Res.* 20, 611–628. doi: 10.1177/009365093020004006

Couper, M. P. (2000). Web surveys: a review of issues and approaches. *Public Opin. Q.* 64, 464–494.

D'Errico, F., and Paciello, M. (2018). Online moral disengagement and hostile emotions in discussions on hosting immigrants. *Internet Res.* 28, 1313–1335. doi: 10.1108/IntR-03-2017-0119

D'Errico, F., and Paciello, M. (2019). Online moral struggles in hosting immigrant's discourses: the underlying role of expressed anger and socio-cognitive processes. *J. Lang. Aggress. Confl.* 7, 182–209. doi: 10.1075/jlac.00024.der

Damasio, A. R. (1998). Emotion in the perspective of an integrated nervous system. *Brain Res. Rev.* 26, 83–86. doi: 10.1016/s0165-0173(97)00064-7

de Vreese, C. H., and Neijens, P. (2016). Measuring media exposure in a changing communications environment. *Commun. Methods Meas.* 10, 69–80. doi: 10.1080/19312458.2016.1150441

Dillman, D. A. (2011). *Mail and Internet Surveys: The Tailored Design Method–2007 Update with New Internet, Visual, and Mixed-Mode Guide*. Hoboken, NJ: John Wiley & Sons.

do Nascimento, J. C. H. B., and da Silva Macedo, M. A. (2016). Structural equation models using partial least squares: an example of the application of SmartPLS§in accounting research. *Rev. Educ. Pesqui. Contabilidade* 10, 282–301. doi: 10.17524/repec.v10i3.1376

Durdyev, S., Ismail, S., Ihtiyar, A., Bakar, N. F. S. A., and Darko, A. (2018). A partial least squares structural equation modeling (PLS-SEM) of barriers to sustainable construction in Malaysia. *J. Clean. Prod.* 204, 564–572. doi: 10.1016/j.jclepro.2018.08.304

Edelstein, M. R. (2002). "Contamination: the invisible built environment," in *Handbook of Environmental Psychology*, eds R. Bechtel and A. Churchman (New York, NY: Wiley), 559–588.

Falk, R. F., and Miller, N. B. (1992). *A Primer for Soft Modeling*. Akron, OH: University of Akron Press.

Finucane, M. L., Alhakami, A., Slovic, P., and Johnson, S. M. (2000). The affect heuristic in judgments of risks and benefits. *J. Behav. Decis. Making* 13, 1–17. doi: 10.1002/(SICI)1099-0771(200001/03)13:1<1::AID-BDM333<3.0.CO;2-S

Fowler, F. J. Jr. (2002). "Survey research methods," in *Applied Social Research Methods Series*, 3rd. Edn. Vol. 1, ed. F. J. Fowler Jr. (Newbury Park, CA: Sage publications).

Fredrickson, B. L. (1998). What good are positive emotions? *Rev. Gen. Psychol.* 2, 300–319. doi: 10.1037/1089-2680.2.3.300

Frewer, L. J., Miles, S., and Marsh, R. (2002). The media and genetically modified foods: evidence in support of social amplification of risk. *Risk Anal. Int. J.* 22, 701–711. doi: 10.1111/0272-4332.00062

Garfin, D. R., Silver, R. C., and Holman, E. A. (2020). The novel coronavirus (COVID-2019) outbreak: amplification of public health consequences by media exposure. *Health Psychol.* 39, 355–357. doi: 10.1037/hea0000875

Gefen, D., Rigdon, E. E., and Straub, D. (2011). Editor's comments: an update and extension to SEM guidelines for administrative and social science research. *MIS Q.* 35, iii–xiv. doi: 10.2307/23044042

Gefen, D., Straub, D., and Boudreau, M.-C. (2000). Structural equation modeling and regression: guidelines for research practice. *Commun. Assoc. Inform. Syst.* 4, 1–78. doi: 10.17705/1CAIS.00407

Glynn, C. J., and Ostman, R. E. (1988). Public opinion about public opinion. *J. Q.* 65, 299–306. doi: 10.1177/107769908806500206

Greenwald, A. G., and Leavitt, C. (1984). Audience involvement in advertising: four levels. *J. Consum. Res.* 11, 581–592. doi: 10.1086/208994

Hair, J. F., Black, B., Babin, B., Anderson, R. E., and Tatham, R. L. (1992). *Multivariate Data Analysis*. New York, NY: Macmillan.

Hair, J. F., Risher, J. J., Sarstedt, M., and Ringle, C. M. (2019). When to use and how to report the results of PLS-SEM. *Eur. Bus. Rev.* 31, 2–24. doi: 10.1108/ebr-11-2018-0203

Harper, C. A., Satchell, L. P., Fido, D., and Latzman, R. D. (2020). Functional fear predicts public health compliance in the COVID-19 pandemic. *Int. J. Ment. Health Addict.* 19, 1875–1888. doi: 10.1007/s11469-020-00281-5

Henrion, M., and Fischhoff, B. (1986). Uncertainty assessment in the estimation of physical constants. *Am. J. Phys.* 54, 791–798. doi: 10.1119/1.14447

Hirschburg, P. L., Dillman, D. A., and Ball-Rokeach, S. J. (1986). "Media system dependency theory: responses to the eruption of Mount St. Helens," in *Media, Audience, and Social Structure*, eds S. J. Ball-Rokeach and M. G. Cantor (Beverly Hills, CA: Sage), 117–126.

Holman, E. A., Garfin, D. R., and Silver, R. C. (2014). Media's role in broadcasting acute stress following the Boston Marathon bombings. *Proc. Natl. Acad. Sci. U.S.A.* 111, 93–98. doi: 10.1073/pnas.1316265110

Hultman, M., Robson, M. J., and Katsikeas, C. S. (2009). Export product strategy fit and performance: an empirical investigation. *J. Int. Mark.* 17, 1–23. doi: 10.1509/jimk.17.4.1

Huynh, T. L. (2020). The COVID-19 risk perception: a survey on socioeconomics and media attention. *Econ. Bull.* 40, 758–764. doi: 10.1016/j.dib.2020.105530

Jackob, N. G. E. (2010). No alternatives? The relationship between perceived media dependency, use of alternative information sources, and general trust in mass media. *Int. J. Commun.* 4, 589–606.

Janz, N. K., and Becker, M. H. (1984). The health belief model: a decade later. *Health Educ. Q.* 11, 1–47. doi: 10.1177/109019818401100101

Jones, N. M., Thompson, R. R., Schetter, C. D., and Silver, R. C. (2017). Distress and rumor exposure on social media during a campus lockdown. *Proc. Natl. Acad. Sci. U.S.A.* 114, 11663–11668. doi: 10.1073/pnas.1708518114

Karasneh, R., Al-Azzam, S., Muflih, S., Soudah, O., Hawamdeh, S., and Khader, Y. (2020). Media's effect on shaping knowledge, awareness risk perceptions and communication practices of pandemic COVID-19 among pharmacists. *Res. Soc. Adm. Pharm.* 17, 1897–1902. doi: 10.1016/j.sapharm.2020.04.027

Kasperson, R. E., Renn, O., Slovic, P., Brown, H. S., Emel, J., Goble, R., et al. (1988). The social amplification of risk: a conceptual framework. *Risk Anal.* 8, 177–187. doi: P10.1111/j.1539-6924.1988.tb01168.x

Khan, G. F., Sarstedt, M., Shiau, W.-L., Hair, J. F., Ringle, C. M., and Fritze, M. P. (2019). Methodological research on partial least squares structural equation modeling (PLS-SEM): an analysis based on social network approaches. *Internet Res.* 29, 407–429. doi: 10.1108/intr-12-2017-0509

Kittipongvises, S., and Mino, T. (2013). The influence of psychological factors on global climate change perceptions held by the rural citizens of Thailand. *Ecopsychology* 5, 126–135. doi: 10.1089/eco.2012.0080

Lachlan, K. A., Spence, P. R., Lin, X., Najarian, K., and Del Greco, M. (2016). Social media and crisis management: CERC, search strategies, and Twitter content. *Comput. Hum. Behav.* 54, 647–652. doi: 10.1016/j.chb.2015.05.027

Lau, J., Yang, X., Tsui, H., and Kim, J. (2003). Monitoring community responses to the SARS epidemic in Hong Kong: from day 10 to day 62. *J. Epidemiol. Commun. Health* 57, 864–870. doi: 10.1136/jech.57.11.864

Lau, J. T., Kim, J. H., Tsui, H., and Griffiths, S. (2007). Perceptions related to human avian influenza and their associations with anticipated psychological and behavioral responses at the onset of outbreak in the Hong Kong Chinese general population. *Am. J. Infect. Control* 35, 38–49. doi: 10.1016/j.ajic.2006.07.010

Lau, J. T., Yang, X., Tsui, H., and Pang, E. (2004). SARS related preventive and risk behaviours practised by Hong Kong-mainland China cross border travellers during the outbreak of the SARS epidemic in Hong Kong. *J. Epidemiol. Commun. Health* 58, 988–996. doi: 10.1136/jech.2003.017483

Lei, J., Rao, Y., Li, Q., Quan, X., and Wenyin, L. (2014). Towards building a social emotion detection system for online news. *Future Gener. Comput. Syst.* 37, 438–448. doi: 10.1016/j.future.2013.09.024

Leiserowitz, A. (2006). Climate change risk perception and policy preferences: the role of affect, imagery, and values. *Clim. Change* 77, 45–72. doi: 10.1007/s10584-006-9059-9

Leppin, A., and Aro, A. R. (2009). Risk perceptions related to SARS and avian influenza: theoretical foundations of current empirical research. *Int. J. Behav. Med.* 16, 7–29. doi: 10.1007/s12529-008-9002-8

Lerner, J. S., Gonzalez, R. M., Small, D. A., and Fischhoff, B. (2003). Effects of fear and anger on perceived risks of terrorism: a national field experiment. *Psychol. Sci.* 14, 144–150. doi: 10.1111/1467-9280.01433

Lerner, J. S., and Keltner, D. (2001). Fear, anger, and risk. *J. Pers. Soc. Psychol.* 81, 146–159. doi: 10.1037/0022-3514.81.1.146

Leung, G., Quah, S., Ho, L., Ho, S., Hedley, A., Lee, H., et al. (2004). A tale of two cities: community psychobehavioral surveillance in Hong Kong and Singapore during the severe acute respiratory syndrome epidemic. *Infect. Control Hosp. Epidemiol.* 25, 1033–1041. doi: 10.1086/502340

Leung, G. M., Ho, L.-M., Chan, S. K. K., Ho, S.-Y., Bacon-Shone, J., Choy, R. Y. L., et al. (2005). Longitudinal assessment of community psychobehavioral responses during and after the 2003 outbreak of severe acute respiratory syndrome in Hong Kong. *Clin. Infect. Dis.* 40, 1713–1720. doi: 10.1086/429923

Leung, G. M., Lam, T.-H., Ho, L.-M., Ho, S.-Y., Chan, B. H. Y., Wong, I. O. L., et al. (2003). The impact of community psychological responses on outbreak control for severe acute respiratory syndrome in Hong Kong. *J. Epidemiol. Commun. Health* 57, 857–863. doi: 10.1136/jech.57.11.857

Lindell, M. K., and Whitney, D. J. (2001). Accounting for common method variance in cross-sectional research designs. *J. Appl. Psychol.* 86, 114–121.

Lion, R., Meertens, R. M., and Bot, I. (2002). Priorities in information desire about unknown risks. *Risk Anal. Int. J.* 22, 765–776. doi: 10.1111/0272-4332.00067

Loewenstein, G. F., Weber, E. U., Hsee, C. K., and Welch, N. (2001). Risk as feelings. *Psychol. Bull.* 127, 267–286. doi: 10.1037/0033-2909.127.2.267

Loges, W. E. (1994). Canaries in the coal mine: perceptions of threat and media system dependency relations. *Commun. Res.* 21, 5–23. doi: 10.1177/009365094021001002

Lundgren, R. E., and McMakin, A. H. (2018). *Risk Communication: A Handbook for Communicating Environmental, Safety, and Health Risks*. Columbus, OH: Battelle Press.

Majeed, S., and Ramkissoon, H. (2020). Health, wellness, and place attachment during and post health pandemics. *Front. Psychol.* 11:573220. doi: 10.3389/fpsyg.2020.573220

Malecki, K., Keating, J. A., and Safdar, N. (2020). Crisis communication and public perception of COVID-19 risk in the era of social media. *Clin. Infect. Dis.* 72, 697–702. doi: 10.1093/cid/ciaa758

Mary, D., and Wildavsky, A. (1982). *Risk and Culture: An Essay on the Selection of Technological and Environmental Dangers*. Berkeley, CA: University of California Press.

Mei-Ling, H. S. U., and Yie-Jing, Y. A. N. G. (2011). Risk communication of global warming: lay perceptions and actions in Taiwan. *Commun. Soc.* 15, 71–104.

Mejia, C. R., Ticona, D., Rodriguez-Alarcon, J. F., Campos-Urbina, A. M., Catay-Medina, J. B., Porta-Quinto, T., et al. (2020). The media and their informative role in the face of the coronavirus disease 2019 (COVID19): validation of fear perception and magnitude of the issue (MEDCOVID-19). *Electron J. Gen. Med.* 17:em239. doi: 10.29333/ejgm/7946

Mertens, G., Gerritsen, L., Duijndam, S., Salemink, E., and Engelhard, I. M. (2020). Fear of the coronavirus (COVID-19): predictors in an online study conducted in March 2020. *J. Anxiety Disord.* 74:102258. doi: 10.1016/j.janxdis.2020.102258

Mikels, J. A., Fredrickson, B. L., Larkin, G. R., Lindberg, C. M., Maglio, S. J., and Reuter-Lorenz, P. A. (2005). Emotional category data on images from the international affective picture system. *Behav. Res. Methods* 37, 626–630. doi: 10.3758/BF03192732

Morgan, M. G., Slovic, P., Nair, I., Geisler, D., MacGregor, D., Fischhoff, B., et al. (1985). Powerline frequency electric and magnetic fields: a pilot study of risk perception. *Risk Anal.* 5, 139–149. doi: 10.1111/j.1539-6924.1985.tb00161.x

Morton, T. A., and Duck, J. M. (2001). Communication and health beliefs: mass and interpersonal influences on perceptions of risk to self and others. *Commun. Res.* 28, 602–626. doi: 10.1177/009365001028005002

Muñiz, C. (2011). Búsqueda de información durante tiempos de crisis. Efectos de la comunicación interpersonal y masiva en la percepción de riesgo personal ante la gripe AH1N1. *Rev. Cien. Soc.* 17, 9–21.

Muñiz, C. (2020). Media system dependency and change in risk perception during the COVID-19 pandemic. *Trípodos* 1, 11–26.

Nerb, J., and Spada, H. (2001). Evaluation of environmental problems: a coherence model of cognition and emotion. *Cogn. Emot.* 15, 521–551. doi: 10.1080/02699930126254

Neuwirth, K., Dunwoody, S., and Griffin, R. J. (2000). Protection motivation and risk communication. *Risk Anal.* 20, 721–734. doi: 10.1111/0272-4332.205065

Ning, L., Niu, J., Bi, X., Yang, C., Liu, Z., Wu, Q., et al. (2020). The impacts of knowledge, risk perception, emotion and information on citizens' protective behaviors during the outbreak of COVID-19: a cross-sectional study in China. *BMC Public Health* 20:1751. doi: 10.1186/s12889-020-09892-y

Nyhan, B., and Reifler, J. (2010). When corrections fail: the persistence of political misperceptions. *Polit. Behav.* 32, 303–330. doi: 10.1007/s11109-010-9112-2

O'Neill, S. J., and Hulme, M. (2009). An iconic approach for representing climate change. *Glob. Environ. Change* 19, 402–410. doi: 10.1016/j.gloenvcha.2009.07.004

Ojala, M. (2012). Hope and climate change: the importance of hope for environmental engagement among young people. *Environ. Educ. Res.* 18, 625–642. doi: 10.1080/13504622.2011.637157

Paciello, M., D'Errico, F., Saleri, G., and Lamponi, E. (2021). Online sexist meme and its effects on moral and emotional processes in social media. *Comput. Hum. Behav.* 116:106655. doi: 10.1016/j.chb.2020.106655

Peng, C.-T., Wu, T.-Y., Chen, Y., and Atkin, D. J. (2019). Comparing and modeling via social media: the social influences of fitspiration on male instagram users' work out intention. *Comput. Hum. Behav.* 99, 156–167.

Peters, E., and Slovic, P. (2000). The springs of action: affective and analytical information processing in choice. *Pers. Soc. Psychol. Bull.* 26, 1465–1475. doi: 10.1177/01461672002612002

Plutchik, R. (1984). "Emotions: a general psychoevolutionary theory," in *Approaches to Emotion*, Vol. 1984, eds K. R. Scherer and P. Ekman (Hillsdale, NJ: Lawrence Erlbaum Associates), 197–219. doi: 10.1177/1937586717705085

Qiao, S., Tam, C. C., and Li, X. (2021). Risk exposures, risk perceptions, negative attitudes toward general vaccination, and COVID-19 vaccine acceptance among college students in South Carolina. *Am. J. Health Promot.* 36, 175–179. doi: 10.1177/08901171211028407

Quah, S. R., and Hin-Peng, L. (2004). Crisis prevention and management during SARS outbreak, Singapore. *Emerg. Infect. Dis.* 10, 364–368. doi: 10.3201/eid1002.030418

Quinn, S. C., Kumar, S., Freimuth, V. S., Kidwell, K., and Musa, D. (2009). Public willingness to take a vaccine or drug under emergency use authorization during the 2009 H1N1 pandemic. *Biosecur. Bioterror.* 7, 275–290. doi: 10.1089/bsp.2009.0041

Ramkissoon, H. (2020). COVID-19 Place confinement, pro-social, pro-environmental behaviors, and residents' wellbeing: a new conceptual framework. *Front. Psychol.* 11:2248. doi: 10.3389/fpsyg.2020.02248

Ramkissoon, H. (2021a). Place affect interventions during and after the COVID-19 pandemic. *Front. Psychol.* 12:726685. doi: 10.3389/fpsyg.2021.726685

Ramkissoon, H. (2021b). Social bonding and public trust/distrust in COVID-19 vaccines. *Sustainability* 13:10248.

Ramkissoon, H. R., and Smith, L. D. G. (2014). The relationship between environmental worldviews, emotions and personal efficacy in climate change. *Int. J. Arts Sci.* 7, 93–109.

Rogers, E. M. (2000). Reflections on news event diffusion research. *J. Mass Commun. Q.* 77, 561–576. doi: 10.1177/107769900007700307

Rogers, R. W., and Prentice-Dunn, S. (1997). Protection motivation theory. *Plenum.*

Rubin, G. J., Amlôt, R., Page, L., and Wessely, S. (2009). Public perceptions, anxiety, and behaviour change in relation to the swine flu outbreak: cross sectional telephone survey. *BMJ* 339:b2651. doi: 10.1136/bmj.b2651

Russell, J. A. (1980). A circumplex model of affect. *J. Pers. Soc. Psychol.* 39, 1161–1178. doi: 10.1037/h0077714

Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychol. Rev.* 110, 145–172. doi: 10.1037/0033-295x.110.1.145

Russell, J. A., and Barrett, L. F. (1999). Core affect, prototypical emotional episodes, and other things called emotion: dissecting the elephant. *J. Pers. Soc. Psychol.* 76, 805–819. doi: 10.1037//0022-3514.76.5.805

Scherer, K. R. (2005). What are emotions? And how can they be measured? *Soc. Sci. Inform.* 44, 695–729. doi: 10.1177/0539018405058216

Shiau, W.-L., and Luo, M. M. (2012). Factors affecting online group buying intention and satisfaction: a social exchange theory perspective. *Comput. Hum. Behav.* 28, 2431–2444. doi: 10.1016/j.chb.2012.07.030

Shiau, W.-L., Sarstedt, M., and Hair, J. F. (2019). Internet research using partial least squares structural equation modeling (PLS-SEM). *Internet Res.* 29, 398–406. doi: 10.1108/intr-10-2018-0447

Skarlatidou, A., Cheng, T., and Haklay, M. (2012). What do lay people want to know about the disposal of nuclear waste? A mental model approach to the design and development of an online risk communication. *Risk Anal. Int. J.* 32, 1496–1511. doi: 10.1111/j.1539-6924.2011.01773.x

Singer, E., and Endreny, P. M. (1993). *Reporting on Risk: How the Mass Media Portray Accidents, Diseases, Other Hazards*. New York, NY: Russell Sage Foundation.

Slovic, P. (1987). Perception of risk. *Science* 236, 280–285.

Slovic, P. (1999). Trust, emotion, sex, politics, and science: surveying the risk-assessment battlefield. *Risk Anal.* 19, 689–701. doi: 10.1023/A:1007041821623

Slovic, P., Finucane, M., Peters, E., MacGregor, D. G., Gilovich, T., Griffin, D., et al. (2002). *Intuitive Judgment: Heuristics and Biases*. Cambridge: University Press Cambridge.

Slovic, P., Finucane, M. L., Peters, E., and MacGregor, D. G. (2004). Risk as analysis and risk as feelings: some thoughts about affect, reason, risk, and rationality. *Risk Anal.* 24, 311–322. doi: 10.1111/j.0272-4332.2004.00433.x

Slovic, P., Fischhoff, B., and Lichtenstein, S. (1980). "Facts and fears: understanding perceived risk," in *Societal Risk Assessment*, eds R. C. Schwing and W. A. Albers (Boston, MA: Springer), 181–216.

Slovic, P. E. (2016). *The Perception of Risk*. Abingdon: Routledge. doi: 10.4324/9781315661773

Smith, N., and Leiserowitz, A. (2012). The rise of global warming skepticism: exploring affective image associations in the United States over time. *Risk Anal.* 32, 1021–1032. doi: 10.1111/j.1539-6924.2012.01801.x

Smith, N., and Leiserowitz, A. (2014). The role of emotion in global warming policy support and opposition. *Risk Anal.* 34, 937–948. doi: 10.1111/risa.12140

Snyder, L. B., and Rouse, R. A. (1995). The media can have more than an impersonal impact: the case of AIDS risk perceptions and behavior. *Health Commun.* 7, 125–145. doi: 10.1207/s15327027hc0702_3

Stevenson, R. A., Mikels, J. A., and James, T. W. (2007). Characterization of the affective norms for English words by discrete emotional categories. *Behav. Res. Methods* 39, 1020–1024. doi: 10.3758/BF03192999

Stúrmer, S., and Simon, B. (2004). Collective action: towards a dual-pathway model. *Eur. Rev. Soc. Psychol.* 15, 59–99. doi: 10.1080/10463280340000117

Taha, S., Matheson, K., Cronin, T., and Anisman, H. (2014). Intolerance of uncertainty, appraisals, coping, and anxiety: the case of the 2009 H 1 N 1 pandemic. *Br. J. Health Psychol.* 19, 592–605. doi: 10.1111/bjhp.12058

Taha, S. A., Matheson, K., and Anisman, H. (2014). H1N1 was not all that scary: uncertainty and stressor appraisals predict anxiety related to a coming viral threat. *Stress Health* 30, 149–157.

Tang, C. S.-K., and Wong, C.-Y. (2005). Psychosocial factors influencing the practice of preventive behaviors against the severe acute respiratory syndrome among older Chinese in Hong Kong. *J. Aging Health* 17, 490–506. doi: 10.1177/0898264305277966

Tsung-Jen, S. (2017a). How message framing and presentation affect pro-environmental behavioral intentions: a dualpathway model perspective. *Coll. J. Commun. Shih Hsin Univ.* 7, 5–36.

Tsung-Jen, S. (2017b). A public-centered climate change communication: theory and practice. *J. Commun. Cult.* 16, 4–41.

Tucker, W. T., and Ferson, S. (2008). "Introduction. Strategies for risk communication: evolution, evidence, experience," in *Strategies for Risk Communication: Evolution, Evidence, Experience*, eds W. T. Tucker, S. Ferson, A. M. Finkel, and D. Slavin (Hoboken, NJ: Blackwell Publishing), ix–xii. doi: 10.1186/s13054-016-1208-6

Tyler, T. R., and Cook, F. L. (1984). The mass media and judgments of risk: distinguishing impact on personal and societal level judgments. *J. Pers. Soc. Psychol.* 47, 693–708. doi: 10.1037/0022-3514.47.4.693

UK Royal Society (1992). *Risk: Analysis, Perception, Management*. London: Royal Society.

Vaughn, P., and Turner, C. (2016). Decoding via coding: analyzing qualitative text data through thematic coding and survey methodologies. *J. Libr. Adm.* 56, 41–51. doi: 10.1080/01930826.2015.1105035

Wen-Long, X. (2020). *Introduction and Application of Statistical Analysis: SPSS Chinese Version + SmartPLS 3 (PLS_SEM), Vol. 3*. Taipei: GOTOP.

Williams, L. J., Vandenberg, R. J., and Edwards, J. R. (2009). 12 structural equation modeling in management research: a guide for improved analysis. *Acad. Manag. Ann.* 3, 543–604. doi: 10.1080/19416520903065683

Witte, K. (1992). Putting the fear back into fear appeals: the extended parallel process model. *Commun. Monogr.* 59, 329–349. doi: 10.1080/03637759209376276

Witte, K., and Allen, M. (2000). A meta-analysis of fear appeals: implications for effective public health campaigns. *Health Educ. Behav.* 27, 591–615. doi: 10.1177/109019810002700506

Witte, K., Meyer, G., and Martell, D. (2001). *Effective Health Risk Messages: A Step-by-Step Guide*. Thousand Oaks, CA: Sage.

Wixom, B. H., and Watson, H. J. (2001). An empirical investigation of the factors affecting data warehousing success. *MIS Q.* 25, 17–41. doi: 10.2307/3250957

Wong, C.-Y., and Tang, C. S.-K. (2005). Practice of habitual and volitional health behaviors to prevent severe acute respiratory syndrome among Chinese adolescents in Hong Kong. *J. Adolesc. Health* 36, 193–200. doi: 10.1016/j.jadohealth.2004.02.024

Yıldırım, M., and Güler, A. (2020). Factor analysis of the COVID-19 perceived risk scale: a preliminary study. *Death Stud.* 1–8. doi: 10.1080/07481187.2020.1784311 [Epub ahead of print].

Yoo, W. (2019). How risk communication via Facebook and Twitter shapes behavioral intentions: the case of fine dust pollution in South Korea. *J. Health Commun.* 24, 663–673. doi: 10.1080/10810730.2019.1655607

Zheng, D., Luo, Q., and Ritchie, B. W. (2021). Afraid to travel after COVID-19? Self-protection, coping and resilience against pandemic 'travel fear'. *Tour. Manage.* 83:104261. doi: 10.1016/j.tourman.2020.104261

Zhou, F., Yu, T., Du, R., Fan, G., Liu, Y., Liu, Z., et al. (2020). Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. *Lancet* 395, 1054–1062. doi: 10.1016/S0140-6736(20)30566-3

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Li and Zhong. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Psychology

ORIGINAL RESEARCH
published: 01 March 2022
doi: 10.3389/fpsyg.2022.782304

Effects of Social Connectedness on the Sharing of Employee-Created Content

Xueting Zhang[1] *and Man Chen*[2]*

[1]Department of Marketing, School of Business, Hunan University, Changsha, China, [2]Department of Business Administration, School of Economics and Management, China University of Geosciences, Wuhan, China

With the popularity of social network platforms (e.g., Facebook and WeChat), users can easily build social connections with others, create content, and even forward or share content. While previous studies on content sharing shed light on either content creator or receiver, this paper is to investigate whether, when, and how the social connectedness of content creator (i.e., employee) and receiver (i.e., employee's friend) jointly influence the sharing likelihood of receiver. We conducted a field study on the largest social media platform and two experiments in China. Study 1 found that well-connected receivers prefer to share content from well-connected employee, and poorly connected receivers prefer to share content from poorly connected employee, but if the content contains promotional information, well-connected receivers are less likely to share it from the well-connected employee. Studies 2 and 3 confirmed these findings and verified that self-enhancement motivation acts as a mediator. The findings suggest that firm should choose the "right" employees who will send content to their "right" friends and caution about the crowd-out effect of promotional content. We provide new insights into the joint effects of creator and receiver, the moderating role of promotional content, and the mediating role of self-enhancement, which enriches both viral marketing and social media literature.

Keywords: employee-created content, content sharing, social connections, social network, promotional content

OPEN ACCESS

Edited by:
Ellen Anke Schmid, Munich University of the Federal Armed Forces, Germany

Reviewed by:
Peijian Song, Nanjing University, China
Zelin Tong, Hainan University, China

***Correspondence:**
Man Chen
chenman027@126.com

Specialty section:
This article was submitted to Human-Media Interaction, a section of the journal Frontiers in Psychology

Received: *24 September 2021*
Accepted: *21 January 2022*
Published: *01 March 2022*

Citation:
Zhang X and Chen M (2022) Effects of Social Connectedness on the Sharing of Employee-Created Content. Front. Psychol. 13:782304. doi: 10.3389/fpsyg.2022.782304

INTRODUCTION

With the popularity of social network platforms (e.g., Facebook and WeChat), users can easily build social connections with others, create content, and even forward or share content (Peng et al., 2018). These platforms provide new communication tools for firms. Beyond firms' official accounts (e.g., brand page at Facebook, Baker et al., 2016; Meire et al., 2019), more and more firms tend to turn their employees into brand champions and encourage them to create content in their personal accounts, which is called *employee-created content* (ECC; Aggarwal et al., 2012). Employee could act as a seed of viral marketing and brand spokesman (Morhart et al., 2009; Zhang et al., 2022). Employee publicly posts content in his or her account, and then, his or her friends receive this content and further share it with their friends and so forth, leading to a widespread of content in the social network. Weber Shandwick company found that 33% of firms encourage their employees to use social media (Shandwick, 2014). Dell, for example, implemented an employee advocacy program that involves 10% of employees on

social media in 2015. Finally, more than 10,000 employees use social media to post content and drove 45,000 clicks to its website. However, sometimes ECC is not effective while employee's friends do not share the content. A total of 54% of brands (e.g., IBM, Adobe, and AT&T) spread content through employees, while only 8% of users share them (Terpening et al., 2016). Therefore, investigating *whether, when, and why* employees' friends are likely to share content from the employees is critical for the success of viral marketing in social network.

The existing literature has identified several crucial factors that affect receiver's sharing related to network and content characteristics (see **Table 1**). Firstly, existing literature studied various types of content, such as positive/negative and informative/persuasive (see **Table 1** for details). Among these content, promotional content which contains discount, coupon, or freebies acts as an important catalyst for marketers to spread marketing information (Lee et al., 2018). In practice, many firms offer promotional content to promote consumer engagement and sales (Ryu and Feick, 2007). However, content receivers maybe not be likely to share promotional content with others since such content could hurt the social image and lead to a "crowding-out effect" (Verlegh et al., 2013; Lee et al., 2018). A further investigation on the role of promotional content is needed.

Secondly, most studies on content sharing considered either content creator or receiver. For example, for content receivers, some scholars found that well-connected receivers are more likely to share content (Yoganarasimhan, 2011), others expressed that poorly connected receivers are more likely to share content (Toubia and Stephen, 2013). For content creators, Hinz et al. (2011) found that well-connected people can trigger more followers to share, while Watts and Dodds (2007) expressed that well-connected people are less important as initiators of large cascades of referrals. Furthermore, while some studies have considered the effects of tie strength and mutual connections between content creator and receiver on content sharing (Aral and Walker, 2014; Peng et al., 2018; Nanne et al., 2021), they ignored the social connectedness of both content creator and receiver. Based on communication theory, both information source and receiver determine receiver's sharing likelihood (Hovland et al., 1953; Berger, 2014). When people receive information from others, they not only care about the type of content but also concern the characteristics of information source, such as network position, social reputation, and identity (Chang and Wu, 2014; Wang et al., 2021).

To resolve the above research gap, we address the following research questions: (1) Whether employee's friends are likely to share employee-created content in social network? (2) Who (i.e., employee's friends) is more willing to share content from whom (i.e., employee) and why? (3) Does promotional content enhance or inhibit sharing?

We conducted a field data analysis and two lab experiments. Study 1 analyzed the creation and sharing behaviors of 20,715 individuals. The Cox proportional hazard model revealed that well-connected receivers prefer to share content from the well-connected employee, and poorly connected receivers prefer to share content from the poorly connected employee, but if the content contains promotional information, well-connected receivers are less likely to share it from the well-connected employee. Studies 2 and 3 confirmed these findings and verified that self-enhancement motivation acts as a mediator.

These findings make significant contributions. First, we investigated both the social connectedness of content creator and receiver, while previous research focused on either content creator or receiver. In this paper, we show that the sharing likelihood of receiver depends on both the social connections of content creator and receiver. Second, we find the crowd-out effect of promotional content in social network. Existing studies have analyzed the impact of promotional content on sharing with mixed findings (Taecharungroj, 2017; Lee et al., 2018). We further study the moderating roles of promotional content and show that when well-connected friends read content from a well-connected employee, promotional content will reduce their sharing likelihood. Third, based on the self-enhancement theory, we explained and tested why well-connected receivers are more likely to share content from well-connected employees. Compared with previous studies, such as Hinz et al. (2011) and Peng et al. (2018), they did not test the underlying mechanism.

We also provide detailed, practical suggestions for managers. Firms need to select the "right" employees to send content to the "right" friends (Lanz et al., 2019). Therefore, the appropriate seeding strategy depends not only on the content creator but also on the content receiver. We found that the firm should encourage high(low-)-connected employees to send content to their high(low-)-connected friends. Moreover, if the content contains promotional information, the firm should prevent well-connected employees from sending promotional content to well-connected friends. Otherwise, the sharing could be counterproductive.

THEORETICAL BACKGROUND AND HYPOTHESES DEVELOPMENT

Literature Related to Content Sharing

In recent years, scholars have studied various factors affecting content sharing based on different social media platforms, such as Twitter (Toubia and Stephen, 2013), YouTube (Libai et al., 2013), and Facebook (Lee et al., 2018). Specifically, the first research stream focused on content characteristics (see **Table 1**). For example, Baker et al. (2016) classified brand content into positive, negative, and neutral, and found that positive content is more viral. Tellis et al. (2019) found that information-focused and commercial content increase sharing, while emotion-focused content inhibits sharing.

Promotional content is the content that contains any forms of monetary and non-monetary promotional information, such as discount, freebie, and coupon (Lee et al., 2018). It is a type of marketing-related information. In practice, many firms often offer various promotional content to attract consumer participation and enhance sales, such as discounts (Wirtz et al., 2013), cash payments (Exley, 2018), and vouchers (Ryu and Feick, 2007), but if the firm cannot accurately target the promotional content to the "right" people, it will increase the

TABLE 1 | Selected studies on content sharing.

Study	Role of content creator	Measure of content receiver's connectedness	Content type	Method	Test mechanism	Main findings
This study	Employee	Degree centrality of employee's friend	Promotional content	Field data and lab experiments	Yes	Well-connected receivers are more likely to share content from well-connected employees because of self-enhancement. When the content contains promotion, well-connected receivers are less likely to share content from well-connected employees.
Berger and Milkman, 2012	Firm		Positive and negative content	Empirical study and experiments	Yes	Positive content is more viral than negative content, and when the positive or negative content has higher arousal, it will be more viral.
Stephen et al., 2015	Firm		Product-, value-, and brand-related information	Empirical study		Product-related and value-related information has no significant effect on customer's sharing, while general brand-related information has a positive effect on customer sharing.
Taecharungroj, 2017	Firm		Information sharing, and Action-inducing content	Empirical study		The action-inducing content had the highest average number of retweets and favorites.
Tellis et al., 2019	Firm		Information-focused, Emotion-focused, and Commercial content	Empirical study and experiment		Information-focused content has a significantly negative effect on sharing, except in risky contexts; positive emotion-focused content has a positive impact on sharing; Commercial content hurts sharing.
Hinz et al., 2011	Firm	Participants' degree centrality		Field experiment, empirical study		Seeding to well-connected and high-betweenness people exhibit a higher response likelihood.
Toubia and Stephen, 2013	User	User's followers		Field Experiment		Users with a low initial number of followers are more likely to contribute content.
Barasch and Berger, 2014	User	Audience size	Self-presenting content, useful content	Experiments	Yes	Consumers with larger audience sizes prefer to share self-presenting content, whereas consumers with smaller audience sizes prefer to share useful content.

marketing cost. Existing studies found that the impacts of promotional content on sharing likelihood could be positive (Wang et al., 2018), negative (Verlegh et al., 2013), or even uncertain (Hong et al., 2017). On the other hand, promotional content can crowd out the people's image and decrease their sharing intention (Sun et al., 2017). Thus, we will consider whether promotional content is effective.

The second stream centers on social connectedness of users. Well-connected people refer to the people with a high number of connections to others, which is measured by degree centrality (Hinz et al., 2011). The greater the degree centrality, the higher connected the person is in social network. Some scholars indicate that well-connected people (i.e., people with higher degree centrality) are conducive to sharing (Hinz et al., 2011). However, other scholars argue that targeting them is not conducive to sharing (Watts and Dodds, 2007). Besides, most studies considered the social connectedness of either content creator (Barasch and Berger, 2014) or content receiver (Meire et al., 2019). Even though some studies have considered both characteristics of creator and receiver, they addressed network overlap (Peng et al., 2018), tie strength (Tuk et al., 2009), and social distance (Hong et al., 2017) between creators and receivers. In this paper, we consider how social connectedness of both content creators and receivers might influence content sharing.

Mechanism: Self-Enhancement Motivation

Self-enhancement is the fundamental human motivation for WOM communication (Alexandrov et al., 2013). Sundaram et al. (1998) define self-enhancement motivation as people's desire to appear smart in front of others, to improve their image. Alexandrov et al. (2013) indicate that self-enhancement motives arise because people want to project a good image to others by sharing information. In general, this motivation implies that people are willing to share information that makes them look good, establishes uniqueness, or helps them gain good social image (Dubois et al., 2016).

Self-enhancement might relate to the number of connections a user has (Toubia and Stephen, 2013; Barasch and Berger, 2014). In practice, the number of connections a person has on social media platforms can provide a signal of the person's influence or popularity (Chen et al., 2017). Social platforms offer a social stage for people to display themselves, so their sharing on these platforms conventionally should seek to enhance their social status and form good impressions among others. More connections also might arouse such self-enhancement motivations more powerfully, which in turn would affect sharing decisions. For example, people with more social connections might actively share content they receive from certain employees, because such information offers better self-enhancement benefits, as we discuss subsequently.

Receiver's Social Connectedness and Content Sharing

In social networks, social connectedness is usually measured by degree centrality or the number of social connections (Goldenberg et al., 2009). Content receiver with a high-degree centrality is well-connected and possess a larger number of connections with others. When receivers have large number of friends, their peer pressure and status considerations might increase (Toubia and Stephen, 2013). Also, based on the self-enhancement theory, well-connected receivers are more concerned about image than poorly connected receivers (Hennig-Thurau et al., 2004). Therefore, when the well-connected people receive content from employees, they will share content more carefully. On the one hand, when well-connected people share content from the employee, they may be seen as "salesman" reducing their social image (Chatterjee, 2015). On the other hand, when receivers have higher number of social connections, the content they shared hardly meets the tastes of all connections. The more connections, the greater the possibility of false matches. Therefore, we suppose that receivers with high-connectedness are less likely to share content from the employee.

H1: Well-connected (versus poorly connected) receivers are less likely to share employee-created content.

Joint Effect of Employee's and Receiver's Social Connectedness

Employees are considered to be experts on firm's product (Eisend, 2004). Besides, the interpersonal nature of employees makes them appear more authentic (Stockman et al., 2017). Based on self-enhancement theory, people can show their social image through content sharing (Hennig-Thurau et al., 2004). In social networks, well-connected people face greater reputational risks, so they are more cautious in sharing to protect their social image (Barasch and Berger, 2014; Chatterjee, 2015). If the well-connected employees are willing to post content, it means that they are willing to stake their reputation, so the reliability of the content appears greater. Besides, content created by a well-connected employee is perceived as more credible (Stockman et al., 2017). When well-connected friends receive content from well-connected employees, sharing can indicate that they have similar tastes, thereby improving receivers' social image. Therefore, we propose that the well-connected receivers are prone to share content from the well-connected employees.

H2: When the content is created by a well-connected employee, well-connected (versus poorly connected) receivers are more likely to share this content (H2a) because of self-enhancement motivation (H2b).

Moderating Role of Promotional Content

While some scholars found that promotional content can increase customer engagement (Ryu and Feick, 2007; Taecharungroj, 2017), others suggested that promotional content could crowd out people's self-image (Exley, 2018). Because promotional content (i.e., discounts and freebies) can change the nature of interpersonal communication, once the promotional content is shared; receivers might regard their relationship with content creator as a norm of monetary incentive, not as a friendship link (Wentzel et al., 2014).

Sharing promotional content stimulates receivers' self-enhancement motivation, due to their suspicion that the creator benefits from sharing that information, which may diminish the credibility of the shared content. Also, promotional content causes receivers to infer less intrinsic motivation and potentially perceive the incentives as the sole driver of recommendation behavior. As we have noted, compared with poorly connected receivers, well-connected receivers tend to be more careful when they worry their sharing might give bad impression to their friends (Tuk et al., 2009; Xiao et al., 2011). Therefore, when well-connected receivers receive promotional content from well-connected employee, it may arouse self-enhancement considerations and lead to more cautious content sharing.

*H3:*Promotional content attenuates the interactive effect of content receiver's and employee's social connectedness on sharing likelihood of receiver.

OVERVIEW OF STUDIES

Figure 1 presents our research framework. To test it, we conducted a field data analysis and two lab experiments (**Table 2**). In Study 1, the field study is to examine the effectiveness of the receiver's social connectedness on sharing, the interaction between employees' and receivers' social connectedness, and the moderating roles of promotional content. Then, with two experimental studies, we replicate the findings of Study 1 in different contexts (restaurant and e-book) and test the underlying mechanisms. Study 2 provides evidence to support H1, H2a, and H2b. Study 3 confirms H2a and H3. In addition, Study 3 also verifies the crowding-out effect of promotional content.

STUDY 1

Data and Measurement

The field study was implemented in WeChat, one of the largest social network platforms in China. Users of this platform form

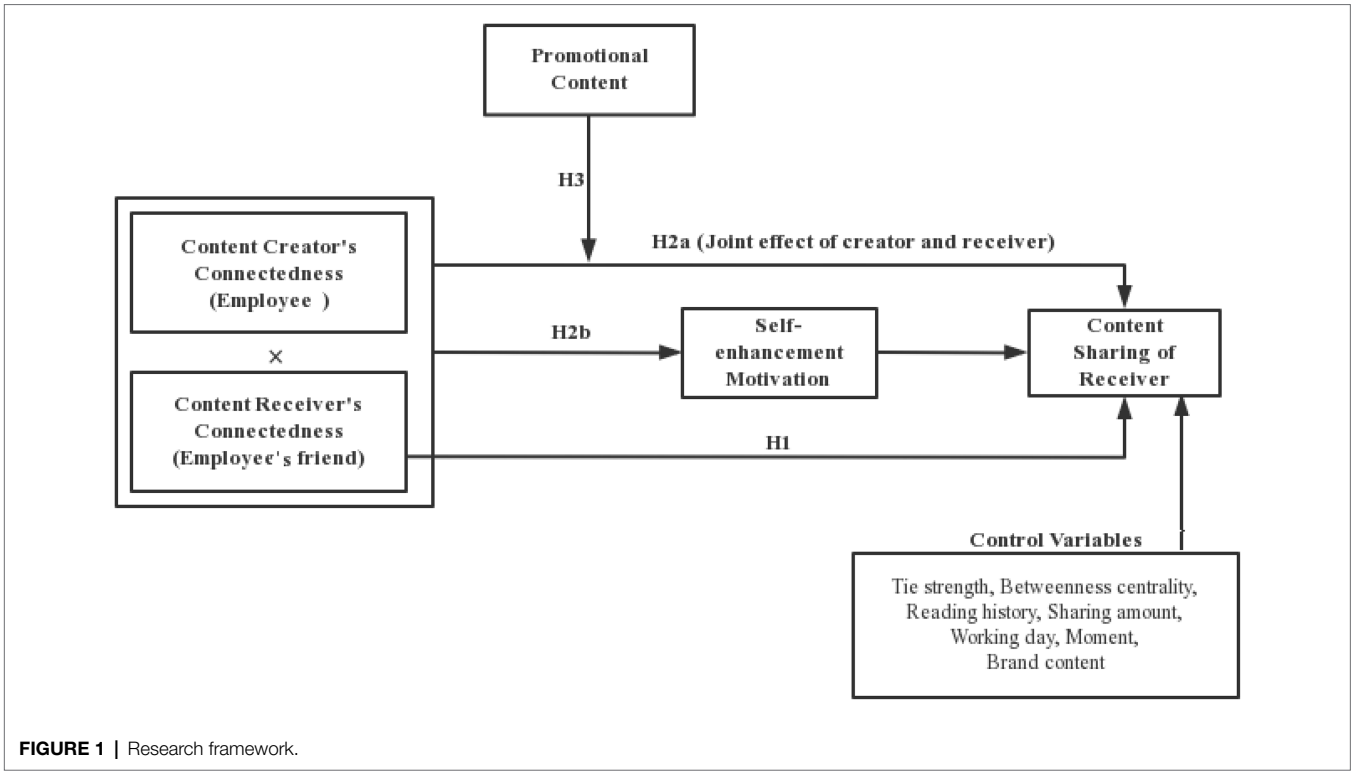


FIGURE 1 | Research framework.

TABLE 2 | Overview of studies.

	Study 1	Study 2	Study 3
Hypothesis	H1, H2a and H3	H1, H2a, and H2b	H2a, H3
Method	Field data analysis	Experiment (2 × 2)	Experiment (2 × 2)
Measurement of social connectedness	Degree centrality	Scenario design	Self-report; 1 item
Content type	promotional, non-promotional		promotional, non-promotional
Test Mechanism		Self-enhancement motivation	Crowd-out effect
Product category	Decoration	e-bookstore	Restaurant
Sample	20,715	139	179

undirected friendship network. We cooperated with an anonymous decoration firm, whose main businesses include home decoration design, building materials selection, and selling furniture. With the popularity of social media in China, this decoration firm often uses WeChat to post marketing-related information in its official account and also encourage employees post-content on their personal WeChat accounts. In WeChat, when an employee creates or posts content, her or his friends could receive and read it, and further decide whether to share or forward it. The decoration firm provided us its employees' content creation and friends' reading and sharing data from June 2017 to November 2017. During this period, the firm let 4,337 employees post 1,270 marketing-related content on their personal accounts.

The dataset contains information about the actual creation, reading, and sharing behaviors of these employees and their friends. In total, 16,368 employees' friends formed a total of 327,736 undirected relationships with these employees. Moreover, these contents were forwarded or shared 184,543 times. The dataset also included the title of each content, the nickname of employees and their friends, sharing time, and the number of reads and shares.

Following existing literature, we measure employee's/receiver's social connectedness as the degree centrality in social network (Hinz et al., 2011). The degree centrality is defined as the number of actors' connections (e.g., friends and fans) that are connected to them (Hinz et al., 2011; Ansari et al., 2018). In our study, if there is a direct friendship between individual i and j, the $a_{ij} = 1$. The degree centrality of individual i is the total number of friends directly connected to i.

$$\text{degree}_{(i)} = \sum_{i \neq j} a_{ij}$$

Next, our dependent variable captures actual content sharing behavior of employees' friends. For example, on Twitter, the content sharing behavior is referred to as "retweet." We coded the dependent variable as 1 when an employee's friend shared the content created from employee, and as 0 when this friend did not share. Then, we coded content that contained promotional information as 1 and 0 otherwise (Lee et al., 2018).

Furthermore, we control the following variables which could influence friend's sharing likelihood. First, the reading time is coded as 1 if it occurred on a workday and 0 for non-work days (i.e., weekends and legal holidays). Second, the title of content containing brand information is coded

as 1 and 0 otherwise. Third, we divided reading moments into two categories, according to whether it was read in the morning. In addition, we also calculated the number of posts that were shared. Following Bond et al. (2012) and Aral and Walker (2014), we controlled the effect of tie strength between content creators (i.e., employee) and receivers (i.e., employee's friends). Tie strength is measured by the frequency of interaction between them before receivers' sharing. In addition, we also calculated the betweenness centrality of content receivers. It captures the extent to people who connect two unconnected groups in the social networks (Hinz et al., 2011). The higher the betweenness centrality, the more likely they act as a network intermediary.

Model Specification

We adopt the hazard modeling method for the survival analysis (Katona et al., 2011; Aral and Walker, 2014). In survival analysis, the occurrence of events is a "failure," and "survival time" refers to the duration of observation, which is generally recorded from the start time of the event to the last recorded time before the event is invalidated or lost. For this study, "failure" refers to whether the receiver forwards or shares the content during the observation time, and "survival time" refers to the period between the post-time of content and the moment the receiver shares it.

We use a Cox proportional hazards model, noting its advantages. First, the model can process censored data and dynamically identify and measure a variety of factors that influence the user's forwarding behaviors. Second, the model considers both the event and the time, which a logistic regression cannot do. Third, the model can include multiple factors that influence information sharing. Fourth, the distribution type for survival time is uncertain, but the model does not depend on a particular distribution. Therefore, we can estimate the impact of the employee's and receiver's degree centrality on the receiver's sharing likelihood while controlling for the moderating effect of promotional content:

$$h\left(t,X_m\right)=h_0\left(t\right)\exp\begin{pmatrix}\beta_1X_{rdegree}+\beta_2X_{edegree}+\beta_3X_{promotional}\\+\beta_4X_{rdegree}*X_{edegree}\\+\beta_5X_{rdegree}*X_{edegree}*X_{promoitonaal}+\beta_6X_{rdegree}*X_{promoitional}+\beta_7*X_{brand}\\+\beta_8X_{readinghistory}+\beta_9X_{moment}\\+\beta_{10}X_{workday}+\beta_{11}X_{sharingamount}\\+\beta_{12}X_{betweenness}+\beta_{13}X_{tie\ strength}\end{pmatrix}$$

Where $X_1,X_2,\ldots X_m$ are concomitant variables and the factors that affect forwarding behavior, which include both static and time-dependent factors; $X_{rdegree}$ denotes the receiver's degree centrality; $X_{edegree}$ represents the employee's degree centrality; are the partial regression coefficients; and $h_0\left(t,X_m\right)$ is the hazard rate for sharing. If the coefficient of the independent variable is positive, the variable X_m accelerates the occurrence of sharing. If the independent variable's coefficient is negative, X_m slows down the occurrence of sharing.

Results and Analysis

Table 3 represents the descriptive statistics. First, it shows that the average number of receivers' social connections is 37; the average number of creators' social connections is 146 (i.e., the degree centrality of employees). The maximum value of social connections is 1,866, and the minimum value is 2. Besides, in **Table 3**, the average variance inflation factor value (VIF) is 1.59, so multicollinearity does not appear to be a concern.

Table 4 shows the results of the Cox proportional hazards model. Model 3 adds the moderating effects of employees' and receivers' degree centrality, beyond the control variable. The results verified H1, we find a negative effect of receiver' degree centrality on content sharing ($\beta_1=-.08$, $p<0.01$), in support H1. Model 4 adds the promotional content variable and analyzes the moderating effects of promotional content on the relationship between receivers' and employees' degree centrality. To support H2a, we find a positive interactive effect of employees' and receivers' degree centrality on sharing ($\beta_1=.05$, $p<0.01$). The moderating effects of promotional content harm the relationship between receivers' and employees' degree centrality ($\beta_2=-.28$, $p<0.01$), in support H3. To depict the interactive effect of the employees' and receivers' degree centrality for two conditions (i.e., promotional vs. non-promotional), we draw three-dimensional diagrams to show the moderating effect of promotion content in **Figure 2**. Specifically, the dark flat indicates the non-promotional condition, and the color flat indicates the promotional condition. This three-dimensional diagram indicates that in the condition of non-promotional content, well-connected friends are more likely to share content from well-connected employees. Whereas in the condition of the promotional content, well-connected friends are less likely to share content from the well-connected employee, which again validates H2a and H3.

STUDY 2

This laboratory experiment has three objectives. First, we repeat the field studies in a more controlled environment to increase validity in the conclusions, and we verify the effect of the receiver's social connectedness on content sharing of ECC. Second, we verify the interactive effects of the employees' and receivers' social connectedness on content sharing. Third, we explore the underlying mechanisms to understand why well-connected receivers are more likely to share content from well-connected employees.

Procedure

The laboratory experiment was conducted at a university in China. We randomly assigned 139 students (52.52% female, $M_{age}=23.52$ years) to a 2×2 between-subjects designs (receiver's social connectedness: high vs. low; sender's social connectedness: high vs. low). To manipulate the receiver's social connectedness, we showed participants a picture that indicated the number

TABLE 3 | Descriptive statistics and correlations.

	Mean	SD	VIF	1	2	3	4	5	6	7	8	9	10
1. Receiver's connectedness	37.08	72.61	3.34	1									
2. Employee's connectedness	146.27	148.87	1.24	0.04	1								
3. Promotional content	0.32	0.47	1.43	−0.06	−0.33	1							
4. Brand	0.30	0.46	1.06	0.01	−0.18	0.05	1						
5. Reading history	1.44	1.18	1.02	−0.01	−0.05	0.13	0.01	1					
6. Working day	0.72	0.45	1.04	0	−0.04	0.1	−0.1	0.03	1				
7. Moment	0.04	0.19	1.00	0	−0.01	0.02	0.02	0	−0.03	1			
8. Sharing amount	992.45	2884.53	1.30	−0.02	−0.13	0.46	−0.07	0.09	0.16	0.02	1		
9. Betweenness centrality	0.01	0.01	3.33	0.84	0.02	−0.02	0	0	0	0	0	1	
10. Tie strength	2.48	4.62	1.12	0	0.31	−0.18	−0.11	−0.03	−0.05	−0.02	−0.10	0	1

VIF, variance inflation factor.

of their friends. The low receiver's social connectedness condition (21) revealed fewer friends than the high receiver's social connectedness condition (101). Another picture showed the number of employees' friends and participants receive the e-book information from the employee. Students are informed that they receive a post related to an e-book from one of their friends who is an employee of an e-book retailer. They are shown a picture and a brief description of this e-book. To check that the manipulation was successful, we asked participants to answer questions about how they perceived the number of friends they had and the employee had (e.g., "What do you think of the number of your connections?" 1 = "not many at all" to 7 = "very many").

In line with Taylor and Todd (1995), we measured participants' sharing intention with four questions ("I intend to share the e-book information to my friends more frequently in the future," "I will try to share the e-book information with my friends," "I will always make an effort to share the e-book information with my friends," and "I intend to share the e-book information with my friends"; 1 = "strongly disagree" to 7 = "strongly agree"; Cronbach's α = 0.95).

Next, we tested the self-enhancement motivation with three items from Dubois et al. (2016): "I shared e-book information so that the message recipient would like me," "I shared e-book information to create a good impression about myself," and "I shared e-book information thinking it would have positive consequences on the message recipient's attitude towards me" (1 = "strongly disagree" to 7 = "strongly agree"; Cronbach's α = 0.96). We also gathered information about participants' gender, age, education, income, and sharing frequency in the social network.

Results and Discussion

A one-way analysis of variance (ANOVA) indicates a significant effect of the receiver's social connectedness [$F(1, 137) = 177.01$, $p < 0.01$], such that participants in the high-connectedness condition indicated that they possessed more connections (M = 6.45, SD = 0.88) than participants in the low-connectedness (M = 3.73, SD = 1.48). The manipulation of employee's social connectedness also was successful [$F(1, 137) = 159.92$, $p < 0.01$], in that participants perceived more connections in the condition of high-connectedness of employee (M = 6.33, SD = 1.06) than in the condition of low-connectedness centrality of employee (M = 3.53, SD = 1.53).

The results of the regression analysis in **Table 5** indicate a significant negative effect of the receiver's connectedness on content sharing ($\beta = -1.64$, $p < 0.05$), in support of H1. Besides, we find a significant positive interaction effect of the receiver's and employee's connectedness ($\beta = 1.31$, $p < 0.01$), in support of H2a. Well-connected receivers are more likely to share content from a well-connected employee. **Figure 3** depicts the joint effect of the employee's and receiver's connectedness on sharing likelihood. In **Figure 3**, the solid line represents the condition of the well-connected employee, and the dotted line represents the condition of the poorly connected employee. **Figure 3** indicates that under the condition of the well-connected employee,

TABLE 4 | Cox proportional hazard regression in Study 1.

Variables		Model 1	Model 2	Model 3	Model 4
		Coef. (S.E.)	Coef. (S.E.)	Coef. (S.E.)	Coef. (S.E.)
Brand		0.26***	0.27***	0.18***	0.15***
		(0.02)	(0.02)	(0.02)	(0.02)
Reading history		0.12***	0.12***	0.11***	0.11***
		(0.00)	(0.00)	(0.00)	(0.00)
Working day		0.04*	0.04*	−0.01	−0.08***
		(0.02)	(0.02)	(0.02)	(0.02)
Moment		−0.25***	−0.25***	−0.26***	−0.30***
		(0.05)	(0.05)	(0.05)	(0.05)
Sharing amount		0.02**	0.02**	−0.01	−0.11***
		(0.01)	(0.01)	(0.01)	(0.01)
Tie strength		−0.62***	−0.62***	−0.41***	−0.32***
		(0.03)	(0.03)	(0.03)	(0.03)
Betweenness centrality		−0.05***	0.04*	0.03	−0.01
		(0.01)	(0.02)	(0.02)	(0.02)
Receiver's connectedness	**H1**		**−0.10*** **	**−0.08*** **	**0.01**
			(0.02)	(0.02)	(0.02)
Employee's connectedness				−0.31***	−0.22***
				(0.01)	(0.01)
Receiver's connectedness × Employee's connectedness	**H2a**			**0.07*** **	**0.05*** **
				(0.01)	(0.01)
Promotional content					0.77***
					(0.02)
Receiver's connectedness × Promotional content					−0.29***
					(0.04)
Employee's connectedness × Receiver's connectedness × Promotional content	**H3**				**−0.28*** **
					(0.05)
Number of failures		10,799	10,799	10,799	10,799
Likelihood ratio χ^2		1,939.85	1,963.35	2,558.20	3,806.59
Log likelihood		−105,602.84	−105,591.09	−105,293.66	−104,669.47

*$p<0.1$, **$p<0.05$, and ***$p<0.01$.

the higher the receiver's social connections, the greater the likelihood of sharing the content, which verified H2a again.

To examine the mediating role of self-enhancement, we use a bootstrapping method to perform the mediation analysis (Hayes, 2017). The significant index of moderated mediation [$b = 1.14$, SE = 0.40, 95% confidence interval (CI) = (0.37, 1.93)] indicates that self-enhancement mediates the interactive effect of employee's and receiver's degree centrality, which supports H2b.

STUDY 3

This laboratory experiment has three further objectives. First, we aim to replicate our finding that well-connected receivers are more likely to share content from well-connected employee (H2a) in another industry setting, namely, a restaurant context. Second, we test the moderating role of promotional content (H3). Third, by examining the potential crowding-out effect of promotional content in this study, we explore why well-connected receivers might be unwilling to share promotional content from well-connected employees.

Procedure

The experiment was conducted at a university in China. We randomly assigned 179 students (44.69% female, $M_{age} = 21.92$ years) to a 2 × 2 between-subjects design (promotional vs. without promotional content; employee's connectedness: high vs. low), with the receiver's social connectedness as a measured variable. To start, the students considered a depiction of the number of the employee's friends. Under the condition of high-degree of employees (151), they have more friends than low-connectedness conditions (21). Next, they were asked to imagine that they received restaurant content from the employee, and we manipulated the promotional content (30% off discount) by providing them with information that either included them or not. To manipulate the senders' connectedness, the procedure was the same as in Study 2. We also used the same items from Study 2 to measure participants' sharing willingness ($\alpha = 0.92$) and self-enhancement motives ($\alpha = 0.91$). We collected the same control variables, reflecting participants' gender, age, education, income, and sharing frequency in the social network.

Results and Discussion

The manipulation of the employee's connectedness is successful [$F(1, 177) = 816.51$, $p < 0.01$], such that compared with the low-connectedness of employee condition (M = 1.96, SD = 0.84), participants in a high-connectedness of employee condition perceived that the employee had more connections (M = 5.64, SD = 0.89).

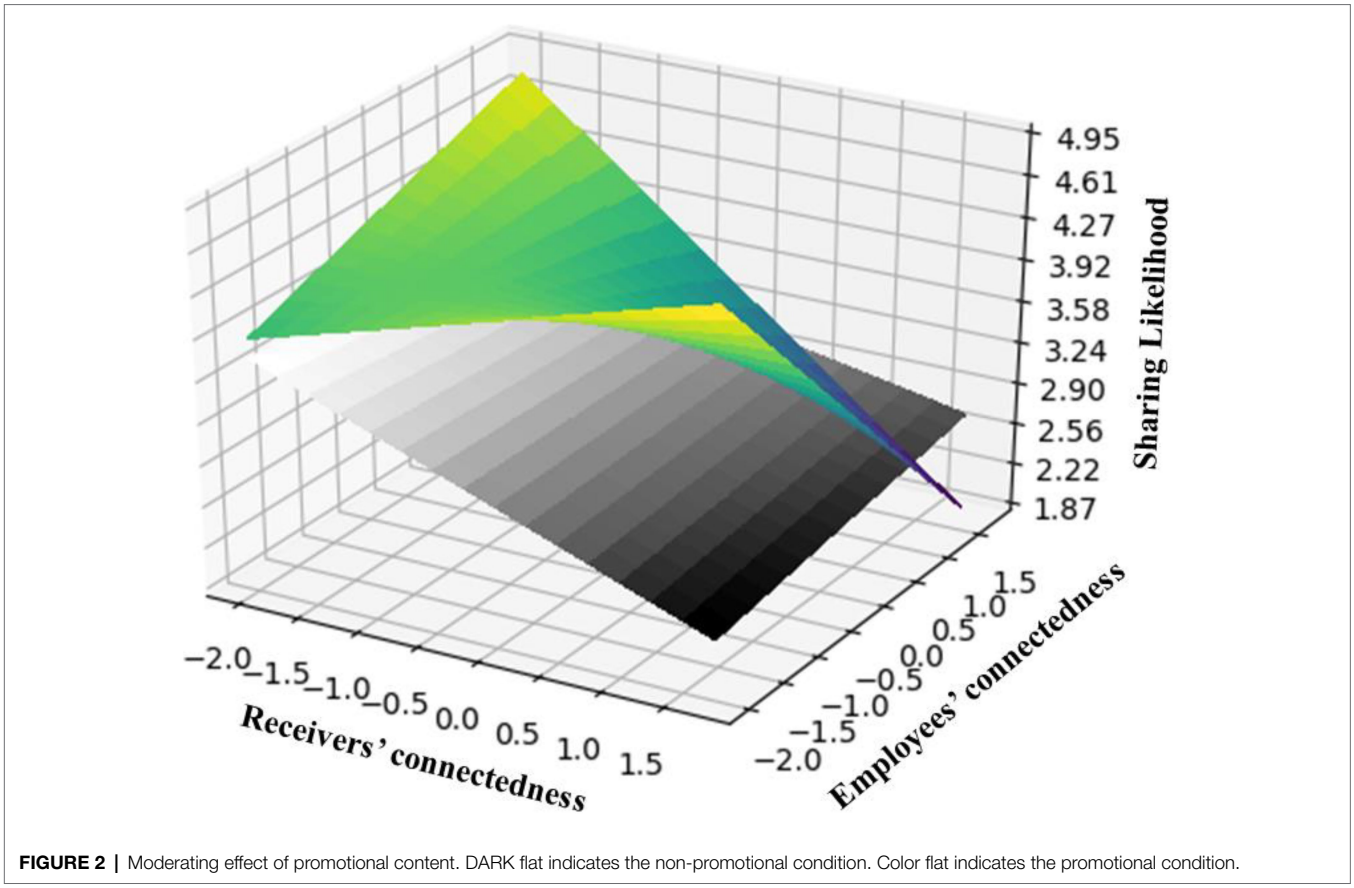


FIGURE 2 | Moderating effect of promotional content. DARK flat indicates the non-promotional condition. Color flat indicates the promotional condition.

The regression analysis indicates a positive interactive effect of the employee's and receiver's connectedness ($\beta = 1.70$, $p < 0.01$). When the information contains promotional, well-connected receivers are less likely to share content from well-connected employees ($\beta = -.63$, $p < 0.05$). We thus find support for both H2a and H3 (**Table 6**). In **Figure 4**, panels a and b illustrate the impact of employees' and receivers' connectedness on content sharing in the promotional and without promotional conditions, respectively. Under the condition of non-promotional content, well-connected receivers (M = 3.85) are more likely to share content from the well-connected employees compared to poorly connected receivers (M = 3.33). Under the condition of promotional content, well-connected receivers (M = 3.89) are less likely to share content from the well-connected employees compared to poorly connected receivers (M = 5.61).

We also test for a crowding-out effect of promotional content. First, we divide the sample into two groups: promotional or without promotional content. Second, we include self-enhancement as a dependent variable and the employee's and receiver's connectedness as independent variables in an ANOVA. **Figure 5** shows the crowd-out effect of promotional content on self-enhancement motivation. Specifically, under the condition of non-promotional content, the self-enhancement motivation of the well-connected receiver (M = 3.46) is higher than that of the poorly connected receiver (M = 2.62) when the content is from the well-connected employee. But under the condition of promotional content, the self-enhancement motivation of the well-connected receiver (M = 3.85) is lower than that of the poorly connected receiver (M = 5.3) when the content is from the well-connected employee. That is to say that promotional information crowds-out people's self-enhancement. In sum, Study 3 verifies H2a and H3 and provides additional support for the crowd-out effect of promotional content.

GENERAL DISCUSSION AND IMPLICATIONS

Although many firms attempt to seed viral campaigns by encouraging well-connected people to disseminate their marketing information, a high failure rate plagues managers and causes controversy in academia (Watts and Peretti, 2007). We explore a new seeding strategy from the perspective of employees. With one field data and two lab experiments, we identify managerially relevant boundary conditions related to social connectedness and promotional content. The studies indicated that well-connected receivers (versus poorly connected) are less likely to share employee-created content, while they prefer to share content from well-connected employees and poorly connected receivers prefer to share content from the poorly connected employee, but if the content contains promotional information (e.g., discount and coupon), well-connected receivers are less likely to share it from the well-connected employee.

TABLE 5 | Moderating role in Study 2.

Variables	Model 1	Model 2	Model 3
	Coef. (S.E.)	Coef. (S.E.)	Coef. (S.E.)
Number of public accounts	−0.07 (0.10)	−0.07 (0.10)	−0.05 (0.1)
Viewing frequency	−0.14 (0.17)	−0.15 (0.17)	−0.12 (0.16)
Sharing frequency	0.33 (0.28)	0.31 (0.28)	0.34 (0.28)
Gender	0.34 (0.24)	0.35 (0.24)	0.32 (0.24)
Age	0.24 (0.26)	0.23 (0.26)	0.32 (0.26)
Education	−0.01 (0.25)	−0.04 (0.25)	−0.09 (0.24)
Income	−0.08 (0.13)	−0.07 (0.13)	−0.11 (0.12)
Receiver's connectedness		0.37 (0.25)	**−1.64**** (0.79)
Employee's connectedness		−0.21 (0.25)	−2.22*** (0.79)
Receiver's connectedness × Employee's connectedness			**1.31***** (0.49)
F	0.79	0.98	1.65
R^2	0.20	0.25	0.34

*$p<0.1$, **$p<0.05$, and ***$p<0.01$.

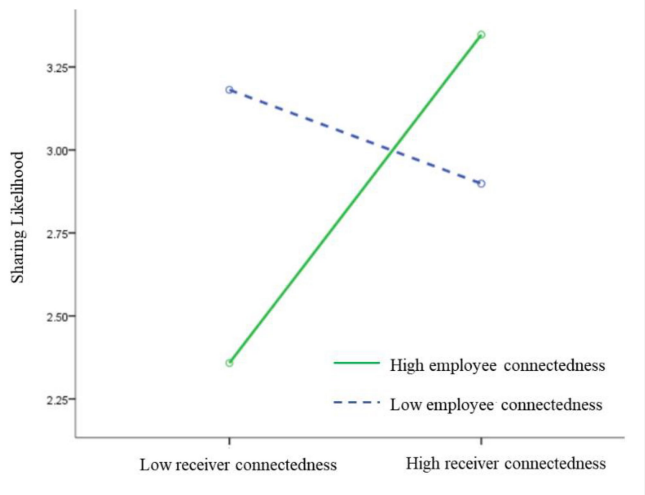


FIGURE 3 | Interaction between employee's and receiver's connectedness in Study 2.

Theoretical Implications

Our research contributes to marketing literature in several ways. First, according to communication theory (Hovland et al., 1953), in addition to personal characteristics, information sources are important determinants of information diffusion. However, extant content marketing literature mostly focuses on personal network characteristics, without considering the role of the information source (Libai et al., 2013). In this paper, we note the social connectedness of information sources as a driver of content sharing, not just the information receiver's social connectedness. Besides, existing findings on the impact of social connectedness on sharing are controversial (Watts and Dodds, 2007; Hinz et al., 2011). Using a field study and lab experiments, we verify that the information source constitutes an important boundary condition. Specifically, well-connected receiver is likely to share information from the well-connected employee, and poorly connected receiver is likely to share information from poorly connected employee.

Second, we explore the underlying mechanism that explains why mismatches in the number of connections between content creators and receivers lead to different effects (Hinz and Spann, 2008). Following self-enhancement theory, we find that well-connected people are more likely to share content from well-connected employees because it stimulates their self-enhancement motivation. As a result, our conclusions enrich both content marketing and social network literature.

Third, this paper provides some new insights into the moderating effects of promotional content. Extant studies of

TABLE 6 | Moderating effect of promotional content in Study 3.

Variables	Model 1	Model 2	Model 3
	Coef. (S.E.)	Coef. (S.E.)	Coef. (S.E.)
Number of public accounts	−0.12 (0.12)	−0.01 (0.1)	−0.03 (0.10)
Viewing frequency	−0.23 (0.19)	−0.05 (0.17)	−0.12 (0.16)
Sharing frequency	0.35** (0.16)	0.27** (0.13)	0.29** (0.13)
Gender	0.19 (0.26)	0.04 (0.22)	0.16 (0.22)
Age	0.04 (0.26)	−0.20 (0.23)	−0.04 (0.22)
Education	−0.20 (0.17)	−0.09 (0.15)	−0.04 (0.14)
Income	−0.05 (0.26)	0.03 (0.23)	−0.13 (0.22)
Receiver's connectedness		**−0.91***** (0.22)	**−1.30** (1.14)
Employee's connectedness		0.61*** (0.23)	−0.55 (0.67)
Promotional content		1.52*** (0.23)	3.62*** (0.66)
Receiver's connectedness × Promotional content			−0.50 (0.61)
Receiver's connectedness × Employee's connectedness			**1.70***** (0.61)
Receiver's connectedness × Employee's connectedness × Promotional content			**−0.63**** (0.28)
R^2	0.22	0.57	0.63
F	1.22	7.86	8.23

*$p<0.1$, **$p<0.05$, and ***$p<0.01$.

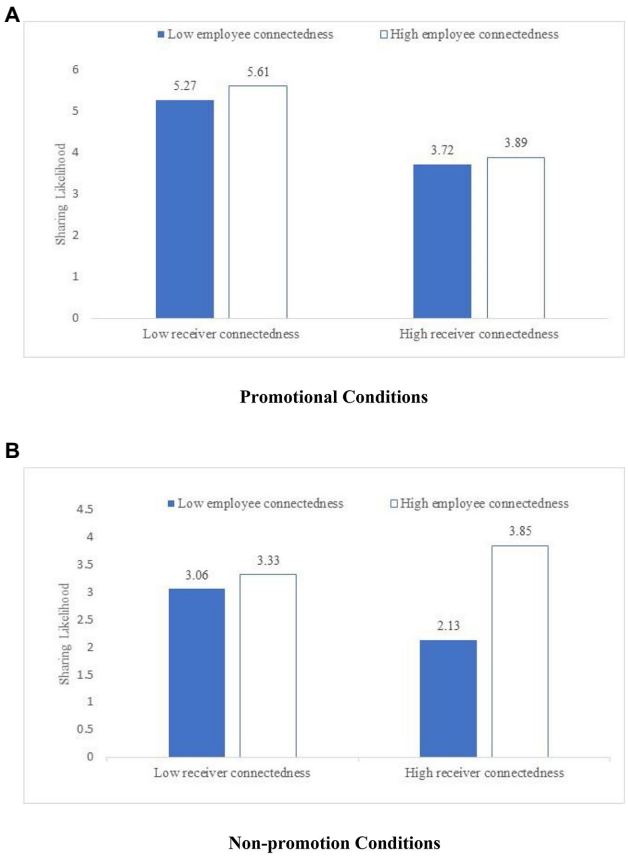


FIGURE 4 | Impact of employee-receiver connectedness on content sharing in Study 3. **(A)** Promotional conditions. **(B)** Non-promotion conditions.

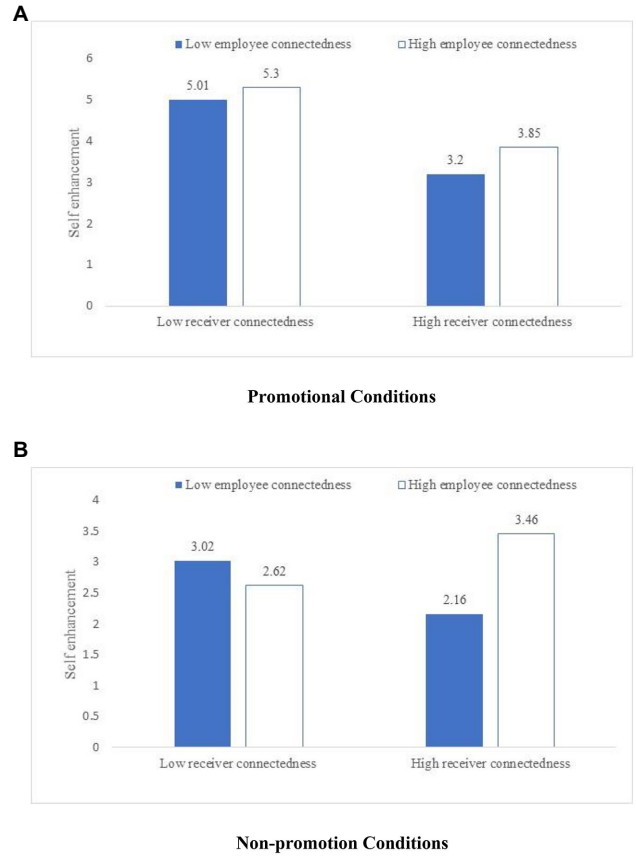


FIGURE 5 | Crowd-out effect of promotional content in Study 3. **(A)** Promotional conditions. **(B)** Non-promotion conditions.

information sharing often consider social network characteristics, and they also report mixed findings of the effectiveness of promotional content for encouraging content sharing (Ryu and Feick, 2007; Hong et al., 2017). Our finding of the interaction between promotional content and social connectedness suggests a new perspective on previous findings. Furthermore, we verify the crowding-out effect of promotional content in social network. Well-connected people who receive promotional content from well-connected employees experience weakened self-enhancement motivations, which decrease their sharing likelihood.

Managerial Implications

This paper provides several implications for managers who want to design effective seeding campaigns and content marketing. First, firms can use sociometric data to improve their effectiveness of marketing campaigns and select the optimal seeding. Particularly, the firm could encourage its employees to post-content on social media, which will facilitate the spread of marketing-related information. However, firms need select the "right" employees who will post-content to the "right" friends (Dost et al., 2018). That is, firms need match the degree centrality of employees and employees' friends to achieve the wide diffusion of content. In particular, well-connected employees should spread content to their well-connected friends.

Second, our research provides firms with more precise promotional strategies in the social networking environment. Especially, we found that the promotional content can be effective and backfire. On the one hand, the firm can leverage the promotional content to stimulate content sharing, especially if well-connected employees spread promotional content to their less-connected friends. Although past research has established that promotional content can increase receivers' sharing intention, it has not taken the impact of social network characteristics into account (Kornish and Li, 2010; Hong et al., 2017). Our research provides suggestions for developing more precise promotional strategies in social networks that can reduce firms' advertising costs.

Third, our findings also caution about the potential backfiring effects of promotional content. When firms encourage employees to spread promotional content, they must balance those employees' and their followers' social connections. Specifically, promotional content can crowd out people's image effects. Therefore, if firms decide to disseminate promotional content, they should avoid allowing well-connected employees to disseminate promotional content to well-connected friends, as promotional content can crowd out friends' image.

Fourth, our research has identified an important psychological mechanism of self-enhancement motivation. Therefore, when

firms disseminate their marketing messages, they should pay attention to the motivation of information receivers, which will help the success of marketing campaigns. Specifically, when well-connected friends receive information from well-connected employee, their self-enhancement motivation will increase, thus improving sharing intention. However, promotional content can undermine friends' self-enhancement motivation. Overall, our findings provide important managerial implications for firms.

Limitations and Further Research

This research offers one of the first empirical demonstrations of a link between promotional content and social network characteristics. However, our data are limited to Chinese social network platform. Further research could verify the generalizability of the conclusions in other cultural and national settings (e.g., Facebook in the United States). Furthermore, we assume the social network is static. Further research could capture the features of network dynamics and also investigate potentially influential factors, such as tie strength and product type (Park et al., 2018).

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

XZ assisted in research, collected and analysed data, and wrote earlier version of the manuscript. MC designed the studies and revised the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This research was supported by the National Natural Science Foundation of China (72072052 and 71772058), Natural Science Foundation of Hunan Province (2020JJ3018), and Key Project from Department of Education of Hunan Province (20A321).

REFERENCES

Aggarwal, R., Gopal, R., Sankaranarayanan, R., and Singh, P. V. (2012). Blog, blogger, and the firm: can negative employee posts lead to positive outcomes? *Inf. Syst. Res.* 23, 306–322. doi: 10.1287/isre.1110.0360

Alexandrov, A., Lilly, B., and Babakus, E. (2013). The effects of social- and self-motives on the intentions to share positive and negative word of mouth. *J. Acad. Mark. Sci.* 41, 531–546. doi: 10.1007/s11747-012-0323-4

Ansari, A., Stahl, F., Heitmann, M., and Bremer, L. (2018). Building a social network for success. *J. Mark. Res.* 55, 321–338. doi: 10.1509/jmr.12.0417

Aral, S., and Walker, D. (2014). Tie strength, embeddedness, and social influence: a large-scale networked experiment. *Manag. Sci.* 60, 1352–1370. doi: 10.1287/mnsc.2014.1936

Baker, A. M., Donthu, N., and Kumar, V. (2016). Investigating how word-of-mouth conversations about brands influence purchase and retransmission intentions. *J. Mark. Res.* 53, 225–239. doi: 10.1509/jmr.14.0099

Barasch, A., and Berger, J. (2014). Broadcasting and narrowcasting: how audience size affects what people share. *J. Mark. Res.* 51, 286–299. doi: 10.1509/jmr.13.0238

Berger, J. (2014). Word of mouth and interpersonal communication: a review and directions for future research. *J. Cons. Psych.* 24, 586–607.

Berger, J., and Milkman, K. L. (2012). What makes online content viral? *J. Mark. Res.* 49, 192–205. doi: 10.1509/jmr.10.0353

Bond, R. M., Fariss, C. J., Jones, J. J., Kramer, A. D. I., Marlow, C., Settle, J. E., et al. (2012). A 61-million-person experiment in social influence and political mobilization. *Nature* 489, 295–298. doi: 10.1038/nature11421

Chang, H. H., and Wu, L. H. (2014). An examination of negative e-WOM adoption: brand commitment as a moderator. *Decis. Support. Syst.* 59, 206–218. doi: 10.1016/j.dss.2013.11.008

Chatterjee, P. (2015). Drivers of new product recommending and referral behaviour on social network sites. *Int. J. Advert.* 30, 77–101. doi: 10.2501/IJA-30-1-077-101

Chen, X., Van Der Lans, R., and Phan, T. Q. (2017). Uncovering the importance of relationship characteristics in social networks: implications for seeding strategies. *J. Mark. Res.* 54, 187–201. doi: 10.1509/jmr.12.0511

Dost, F., Phieler, U., Haenlein, M., and Libai, B. (2018). Seeding as part of the marketing mix: word-of-mouth program interactions for fast-moving consumer goods. *J. Mark.* 83, 62–81. doi: 10.1177/0022242918817000

Dubois, D., Bonezzi, A., and De Angelis, M. (2016). Sharing with friends versus strangers: how interpersonal closeness influences word-of-mouth valence. *J. Mark. Res.* 53, 712–727. doi: 10.1509/jmr.13.0312

Eisend, M. (2004). Is it still worth to be credible? A meta-analysis of temporal patterns of source credibility effects in marketing. *Adv. Consum. Res.* 31, 352–357.

Exley, C. (2018). Incentives for Prosocial behavior: The role of reputations. *Manag. Sci.* 64, 2460–2471. doi: 10.1287/mnsc.2016.2685

Goldenberg, J., Han, S., Lehmann, D. R., and Hong, J. W. (2009). The role of hubs in the adoption process. *J. Mark.* 73, 1–13. doi: 10.1509/jmkg.73.2.1

Hayes, A.F. (2017). *Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach.* New York: Guilford Press.

Hennig-Thurau, T., Gwinner, K. P., Walsh, G., and Gremler, D. D. (2004). Electronic word-of-mouth via consumer-opinion platforms: what motivates consumers to articulate themselves on the internet? *J. Interact. Mark.* 18, 38–52. doi: 10.1002/dir.10073

Hinz, O., Skiera, B., Barrot, C., and Becker, J. U. (2011). Seeding strategies for viral marketing: An empirical comparison. *J. Mark.* 75, 55–71. doi: 10.1509/jm.10.0088

Hinz, O., and Spann, M. (2008). The impact of information diffusion on bidding behavior in secret reserve Price auctions. *Inf. Syst. Res.* 19, 351–368. doi: 10.1287/isre.1080.0190

Hong, Y., Pavlou, P. A., Shi, N., and Wang, K. (2017). On the role of fairness and social distance in designing effective social referral systems. *MIS Q.* 41, 787–809. doi: 10.25300/MISQ/2017/41.3.06

Hovland, C.I., Janis, I.L., and Kelley, H.H. (1953). *Communication and Persuasion.* New Haven: Yale University Press.

Katona, Z., Zubcsek, P. P., and Sarvary, M. (2011). Network effects and personal influences: The diffusion of an online social network. *J. Mark. Res.* 48, 425–443. doi: 10.1509/jmkr.48.3.425

Kornish, L. J., and Li, Q. (2010). Optimal referral bonuses with asymmetric information: firm-offered and interpersonal incentives. *Mark. Sci.* 29, 108–121. doi: 10.1287/mksc.1080.0484

Lanz, A., Goldenberg, J., Shapira, D., and Stahl, F. (2019). Climb or jump: status-based seeding in user-generated content networks. *J. Mark. Res.* 56, 361–378. doi: 10.1177/0022243718824081

Lee, D., Hosanagar, K., and Nair, H. S. (2018). Advertising content and consumer engagement on social media: Evidence from Facebook. *Manag. Sci.* 64, 5105–5131. doi: 10.1287/mnsc.2017.2902

Libai, B., Muller, E., and Peres, R. (2013). Decomposing the value of word-of-mouth seeding programs: acceleration versus expansion. *J. Mark. Res.* 50, 161–176. doi: 10.1509/jmr.11.0305

Meire, M., Hewett, K., Ballings, M., Kumar, V., and Van Den Poel, D. (2019). The role of marketer-generated content in customer engagement marketing. *J. Mark.* 83, 21–42. doi: 10.1177/0022242919873903

Morhart, F. M., Herzog, W., and Tomczak, T. (2009). Brand-specific leadership: turning employees into brand champions.(report). *J. Mark.* 73, 122–142. doi: 10.1509/jmkg.73.5.122

Nanne, A. J., Antheunis, M. L., and Van Noort, G. (2021). The role of facial expression and tie strength in sender presence effects on consumers' brand responses towards visual brand-related user generated content. *Comput. Hum. Behav.* 117, 106628–106628. doi: 10.1016/j.chb.2020.106628

Park, E., Rishika, R., Janakiraman, R., Houston, M. B., and Yoo, B. (2018). Social dollars in online communities: The effect of product, user, and network characteristics. *J. Mark.* 82, 93–114. doi: 10.1509/jm.16.0271

Peng, J., Agarwal, A., Hosanagar, K., and Iyengar, R. (2018). Network overlap and content sharing on social media platforms. *J. Mark. Res.* 55, 571–585. doi: 10.1509/jmr.14.0643

Ryu, G., and Feick, L. (2007). A penny for your thoughts: referral reward programs and referral likelihood. *J. Mark.* 71, 84–94. doi: 10.1509/jmkg.71.1.084

Shandwick, W. (2014). "Employees rising: Seizing the opportunity in employee activism," in: Retrieved December.

Stephen, A.T., Sciandra, M.R., and Inman, J.J. (2015). *The Effects of Content Characteristics on Consumer Engagement with Branded Social Media Content on Facebook*. Cambridge: Marketing Science Institute.

Stockman, S., Van Hoye, G., and Carpentier, M. (2017). The dark side of employee referral bonus programs: potential applicants, awareness of a referral bonus and perceptions of organisational attractiveness. *Appl. Psychol. Int. Rev.* 66, 599–627. doi: 10.1111/apps.12100

Sun, Y., Dong, X., and Mcintyre, S. (2017). Motivation of user-generated content: social connectedness moderates the effects of monetary rewards. *Mark. Sci.* 36, 329–337. doi: 10.1287/mksc.2016.1022

Sundaram, D. S., Mitra, K., and Webster, C. (1998). Word-of-mouth communications: a motivational analysis. *Adv. Consumer Res. Assoc. Consum. Res.* 25, 527–531.

Taecharungroj, V. (2017). Starbucks' marketing communications strategy on Twitter. *J. Mark. Commun.* 23, 552–571. doi: 10.1080/13527266.2016.1138139

Taylor, S., and Todd, P. A. (1995). Understanding information technology usage: a test of competing models. *Inf. Syst. Res.* 6, 144–176. doi: 10.1287/isre.6.2.144

Tellis, G. J., Macinnis, D. J., Tirunillai, S., and Zhang, Y. (2019). What drives virality (sharing) of online digital content? The critical role of information, emotion, and brand prominence. *J. Mark.* 83, 1–20. doi: 10.1177/0022242919841034

Terpening, E., Li, C., and Littleton, A. (2016). Social media employee advocacy: tapping into the power of an engaged social workforce. Available at: https://www.socialdialog.eu/media/Social-Media-Employee-Advocacy-Altimeter.pdf (Accessed December 20, 2021).

Toubia, O., and Stephen, A. T. (2013). Intrinsic vs. image-related utility in social media: why do people contribute content to twitter? *Mark. Sci.* 32, 368–392. doi: 10.1287/mksc.2013.0773

Tuk, M. A., Verlegh, P. W. J., Smidts, A., and Wigboldus, D. H. J. (2009). Sales and sincerity: the role of relational framing in word-of-mouth marketing. *J. Consum. Psychol.* 19, 38–47. doi: 10.1016/j.jcps.2008.12.007

Verlegh, P. W. J., Ryu, G., Tuk, M. A., and Feick, L. (2013). Receiver responses to rewarded referrals: the motive inferences framework. *J. Acad. Mark. Sci.* 41, 669–682. doi: 10.1007/s11747-013-0327-8

Wang, Q., Sun, Y., Zhu, J., and Zhang, X. (2018). The impact of uncertain rewards on customers' recommendation intention in social networks. *Internet Res.* 28, 1029–1054. doi: 10.1108/IntR-03-2017-0116

Wang, W., Shao, T., Yi, Y., Fang, S., Song, J., and Yu, Z. (2021). Subtle signals of status on social network sites: evidence from China. *Front. Psych.* 12, 1–10.

Watts, D. J., and Dodds, P. S. (2007). Influentials, networks, and public opinion formation. *J. Consum. Res.* 34, 441–458. doi: 10.1086/518527

Watts, D. J., and Peretti, J. (2007). Viral marketing for the real world. *Harv. Bus. Rev.* 85, 22–23.

Wentzel, D., Tomczak, T., and Henkel, S. (2014). Can friends also become customers? The impact of employee referral programs on referral likelihood. *J. Serv. Res.* 17, 119–133. doi: 10.1177/1094670513518271

Wirtz, J., Orsingher, C., Chew, P., and Tambyah, S. K. (2013). The role of Metaperception on the effectiveness of referral reward programs. *J. Serv. Res.* 16, 82–98. doi: 10.1177/1094670512462138

Xiao, P., Tang, C. S., and Wirtz, J. (2011). Optimizing referral reward programs under impression management considerations. *Eur. J. Oper. Res.* 215, 730–739. doi: 10.1016/j.ejor.2011.05.042

Yoganarasimhan, H. (2011). Impact of social network structure on content propagation: a study using YouTube data. *Quant. Mark. Econ.* 10, 111–150. doi: 10.1007/s11129-011-9105-4

Zhang, X., Tang, T., Chen, M., and Wang, F. (2022). Effects of employee identity salience on content sharing in social networks. *Mark. Intell. Plan.* 40, 89–104. doi: 10.1108/MIP-09-2021-0304

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Zhang and Chen. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Psychology

ORIGINAL RESEARCH
published: 08 March 2022
doi: 10.3389/fpsyg.2022.754855

Coordinating Mechanisms Are More Important Than Team Processes for Geographically Dispersed Emergency Dispatch and Paramedic Teams

Bjørn Helge Johnsen[1†], Roar Espevik[2,3†], Jarle Eid[2†], Øyvind Østerås[4†], Johan Kolstad Jacobsen[1†] and Guttorm Brattebø[4,5†]*

[1] Department of Psychosocial Science, University of Bergen, Bergen, Norway, [2] Centre for Crisis Psychology, University of Bergen, Bergen, Norway, [3] Swedish Defence University, Stockholm, Sweden, [4] Department of Anaesthesia and Intensive Care, Haukeland University Hospital, Bergen, Norway, [5] Department of Clinical Medicine, University of Bergen, Bergen, Norway

OPEN ACCESS

Edited by:
Oliver Niebuhr, University of Southern Denmark, Denmark

Reviewed by:
Cesar Collazos, University of Cauca, Colombia
Lauren E. Benishek, Johns Hopkins University, United States

***Correspondence:**
Bjørn Helge Johnsen
bjorn.johnsen@uib.no

† These authors have contributed equally to this work

Specialty section:
This article was submitted to Performance Science, a section of the journal Frontiers in Psychology

Received: *07 August 2021*
Accepted: *09 February 2022*
Published: *08 March 2022*

Citation:
Johnsen BH, Espevik R, Eid J, Østerås Ø, Jacobsen JK and Brattebø G (2022) Coordinating Mechanisms Are More Important Than Team Processes for Geographically Dispersed Emergency Dispatch and Paramedic Teams. Front. Psychol. 13:754855. doi: 10.3389/fpsyg.2022.754855

In recent decades there has been an increased emphasis on non-technical skills in medical teams. One promising approach that relates teamwork to medical efficiency is the theory of Shared Mental Models (SMM). The aim of the present study was to investigate the suitability of the Shared Mental Model approach for teamwork between operators in emergency medical communication centers and the first line ambulance personnel in real-life settings. These teams collaborate while working from geographically dispersed positions, which makes them distinct from the kinds of teams examined in most previous research on team effectiveness. A pressing issue is therefore whether current models on co-located teams are valid for medical distributed teams. A total of 240 participants from 80 emergency medical teams participated in the study. A team effectiveness model was proposed based on identified team coordinating mechanisms and the "Big five" team processes. Path analyses showed that SMM was positively associated with team effectiveness (i.e., performance satisfaction and situational awareness) and negatively related to mission complexity. Furthermore, the coordinating mechanisms of SMM and Closed Loop Communication was positively related to "Big five" team scores. However, no effects were found for the "Big five" team processes on effectiveness, which could indicate that the model needs to be adjusted for application to geographically dispersed teams. Possible implications for team training of distributed emergency response teams are discussed.

Keywords: medical first responder teams, shared mental models, team processes, coordinating mechanisms, performance

INTRODUCTION

Emergency medical organization in Norway is centered around emergency medical communication centers (EMCC). The EMCC's are the primary point of contact in case of a critical medical predicament. The personnel at the EMCC consist of registered nurses or paramedics who have completed mandatory training in emergency communication, emergency

medical system knowledge, and the technical skills required for the job. The role of the EMCC is to provide medical advice to the caller based on the Norwegian Index for Emergency Medical Assistance, a decision support tool (Norwegian Medical Association, 2009). This involves conducting triage, dispatching and directing the ambulances, and acting as a communication-hub between the ambulances, general practitioners, and external entities like the local police operational headquarters or fire and rescue units. The EMCC-operator and the ambulance personnel can be considered as constituting a team, since they consist of two or more persons coordinating their activities toward a common goal (Parelta et al., 2018). Given the critical nature of the missions and the need for a rapid response, combined with the complexity of the EMCC operations, there is a constant need to monitor and improve the performance of these frontline services. One promising strategy for augmenting EMCC operations has been to focus on generic and specific non-technical skills (NTS) to improve teamwork. The importance of NTS in health care is documented by Flin and Maran (2004) who outlined a framework for training medical teams. Furthermore, the NTS of team leadership was stressed by the UK General Medical Council (2016), the Academy of Royal Medical Colleges and the UK National Health Service (2010) in their "Medical Leadership Competency Framework."

A promising development in the study of NTS in medical teams is the so-called Shared Mental Model (SMM) approach (Johnsen et al., 2017). The core of the SMM approach holds the view that effective performance during high workload operations relies upon a shared description, understanding and prediction of the occurring events (i.e., SMM; Weld et al., 2016). The shared cognitive construct generates an immediate and internalized understanding of members' coordination, decisions, and actions. Based on a literature review covering 20 years of research on team effectiveness, Salas et al. (2005) proposed a model where SMM, Trust, and Closed-Loop Communication (CLC) acted as coordinating mechanisms in order to develop and/or maintain effective team processes. Based on the literature review, they distilled five team processes called "the Big five in teamwork," which included team leadership, mutual monitoring, team adaptation, team orientation and mutual support behavior. Thus, a relationship between coordinating mechanisms and performance was inferred to be mediated by these five team processes (Salas et al., 2005; Uitdewilligen et al., 2018).

The model proposed by Salas et al. (2005) found empirical support from qualitative and exploratory studies of nursing teams from acute, intensive care, and maternity units (Kalisch et al., 2009). By means of focus groups and probing questions, centered on both coordinating mechanisms as well as team processes, the results indicated that these factors were related to team performance. Another study involving the two coordinating mechanisms of SMM and Trust (McComb et al., 2017) found differences in SMM concerning role responsibilities between physicians and nurses as well as a differential reporting of trust between the two professions. Although the study from McComb et al. (2017) did not test the total SMM model and its relation to team performance, it indicated that there were differences in the use of coordinating mechanisms between nurses and physicians.

In a comparison between two approaches to NTS in health care, Westli et al. (2010) showed that SMM measures explained performance indicators over and beyond the effect of Anesthetists' NTS Behavioral marker system (Fletcher et al., 2003) in acute medical teams. This was also evident in a study of behavioral markers of SMM in high performance trauma teams where leaders of these teams more frequently displayed behavioral markers of SMM compared to less effective teams (Johnsen et al., 2017). However, both studies investigated co-located teams, where the team processes could be observed by all team members. The EMCCs, on the other hand, are characterized by team members solely dependent on verbal coordination while working from separate locations. Sharing information on status, intentions and resource coordination is done exclusively *via* communication technology. It is therefore reasonable to assume that geographical disparity and the dynamic nature of the situation create a challenge for team processes like Mutual Monitoring, Adaptation and Support Behavior.

Modes and ways of communicating have an influence on SMM, and the introduction of new technology heavily influences the flow of information between team members (Meynard and Gilson, 2014). Historically, there has been a change in how teams execute teamwork, which is driven by advances in information technology. Most studies on the development of SMM and its relations to team effectiveness is anchored in an understanding that the relation is motivated by increased interaction between team members, increased communication, and training (Kraiger and Wenzel, 1997; Mjelde et al., 2015). However, little attention has been given to the effect of using only technological means to communicate on the relationship between coordinating mechanisms, team processes and team effectiveness. Both in general and for operational professionals in particular, communication technology has changed teamwork from face-to-face to more virtual interactions where technological assets play an increasingly important role in the command, control, and coordination of distributed emergency response units. The potential for information exchange with technological aids could facilitate, impair, or have a neutral effect on team effectiveness (Curseu et al., 2008; Algesheimer et al., 2011; Meynard and Gilson, 2014). Teams could vary in the degree of reliance on technological tools to communicate, coordinate and execute team processes. Meynard and Gilson (2014) investigated the development of SMM in teams with unknown team members, the interaction of task interdependence, and technological attributes. However, the focus of that study was the development of a new model and not testing the total model proposed by Salas et al. (2005). To our knowledge, no studies on virtual teams have investigated the total effect of both coordinating mechanisms and team processes on team effectiveness in the same analyses using real-life scenarios. Taken together, there is a need for empirical studies of team processes and performance of geographically distributed team members in the Emergency Medical Services (EMS) domain.

The overarching aim of the present study was to test the fitness of the SMM approach on EMS teams where the coordinator is dispersed geographically from the other team members. An evidence-based, theoretical approach (e.g., SMM) could provide

more targeted and effective approaches to the education and training of EMS teams. The suitability of the model could be investigated by using path analysis to explain the separate impact on effectiveness ratings of coordinating mechanisms and team processes. The specific hypotheses are outlined as paths (see **Figure 1**) and listed below:

Hypothesis 1: A direct positive effect of SMM on the measures of performance satisfaction and situational awareness as well as a direct negative effect on perceived mission complexity.

This hypothesis is based on the abundance of studies that has demonstrated an effect of the coordinating mechanism of SMM on team effectiveness (Espevik et al., 2006, 2011a,b; Curseu et al., 2008; Johnson et al., 2011; Meynard and Gilson, 2014; Johnsen et al., 2017).

Hypothesis 2: A direct positive effect of Closed Loop Communication (CLC) on the measures of performance satisfaction and situational awareness as well as a direct negative effect on perceived mission complexity.

CLC has also been shown to affect team output. A link between CLC and performance is found in other domains like firefighter teams (Jouanne et al., 2017). However, in the medical domain, a leadership style involving CLC has been less prominent despite research suggesting that CLC may be vital to success in operational teams (Undre et al., 2006; Härgenstam et al., 2013; Mesmer-Magnus et al., 2017; El-Shafy et al., 2018).

Hypothesis 3: A positive path flowing from SMM through team processes and further on to team effectiveness for the measures of performance satisfaction and situational awareness. Since lower mission complexity represent increased team effectiveness a negative association was hypothesized between team processes and mission complexity in this mediation analyzes.

Hypothesis 4: A positive path flowing from CLC through team processes and further on to team effectiveness for the measures of performance satisfaction and situational awareness. Also, for this mediation analyses lower mission complexity represent increased team effectiveness. Thus, a negative association was hypothesized between team processes and mission complexity.

The pathway from coordinating mechanisms (hypotheses 3 and 4) through team processes is based on Salas et al. (2005), and the assumption that SMM and CLC generate and maintain team processes is well supported in the literature (Johnson-Laird, 1983; Rentsch and Klimoski, 2001; Espevik et al., 2006, 2011b; DeChurch and Mesmer-Magnus, 2010; Johnson et al., 2011; Meynard and Gilson, 2014; Jouanne et al., 2017).

Hypothesis 5: All three coordinating mechanisms, including trust, would be positively inter-correlated and positively correlated with team processes. It could be argued that levels of trust influence coordination within a team, acceptance of mutual monitoring, willingness to support each other, inclination to share information as well as a proclivity to adapt strategies based on input from other team members (McComb et al., 2017).

No specific hypothesis was generated regarding trust and its effect on team output, since trust differs from the other coordinating mechanism by not involving behavior directly related to information sharing in order to alter the cognitive representations of the situation. Instead, trust works more in line with belief systems involving emotional aspects of team membership.

MATERIALS AND METHODS

Participants and Environment

A total of 240 participants constituting 80 teams participated in the study. The teams consisted of two ambulance personnel and one dispatcher in the Bergen EMCC. Bergen is the second largest city in Norway, and Bergen EMCC covers approximately 450,000 people, with about 58,000 ambulance missions a year. Fifty-six percent of the operators reported less than one year of EMCC experience, 40% had between one and 12 years of experience, and 4% reported more than 12 years of experience.

Measures

SMMs, trust, performance satisfaction, situational awareness, and mission complexity were measured using a five-item, single-page, paper-based, self-report questionnaire. Each construct was measured using a single item 100 mm visual analog scale (VAS; McCormack et al., 1988). The VAS was constructed as a solid line with anchor points representing opposite descriptors (e.g., very low to very high).

Item details are as follows:

Shared Mental Models. Shared mental models. SMMs were rated with the question "To what degree did the EMCC-operator and ambulance personnel create a shared understanding/SMM of the mission (team members updated each other on aims, situations, priorities, and internal/external needs)?"

Trust. Trust was measured with the item: "To what degree did the EMCC-operator and the ambulance personnel trust each other in order to ensure that the mission was completed in the best possible way? Mark your assessment of trust between the EMCC-operator and ambulance personnel."

Performance satisfaction. Subjective ratings are thought to be better measures of performance than objective measures since objective measures fail to account for the impact of non-controllable third variables (Nicholls et al., 2012). Sports psychology research has advocated for the use of subjective performance measures given the importance of emotion in performance appraisal (i.e., feelings of satisfaction; Nicholls et al., 2012). Therefore, performance satisfaction was measured using the item: "How satisfied were you with the performance of the team?"

Situational awareness. Operators' perception of the team's overall situational awareness was measured using the

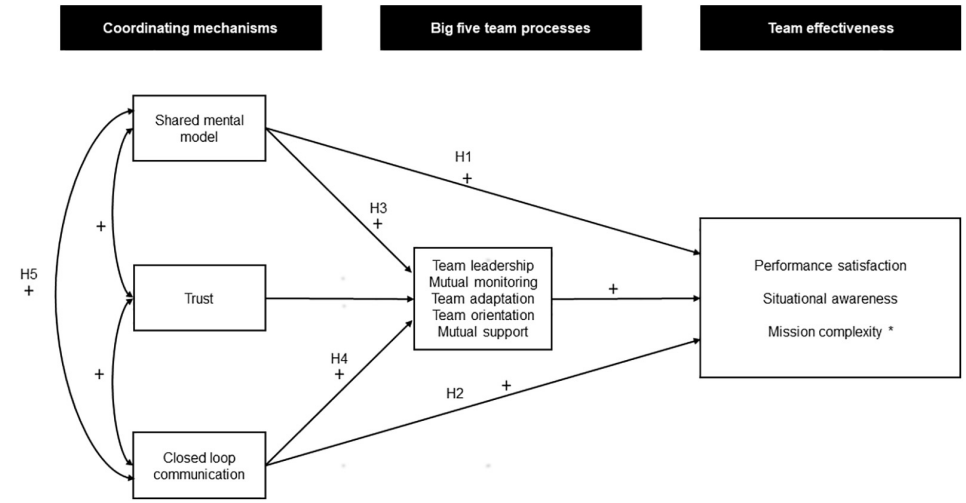


FIGURE 1 | The proposed model of the study. The model is separated for the three coordinating mechanisms of trust, shared mental model and closed loop communication. The Big five team processes are expressed as a composite score of team leadership, mutual monitoring, team adaptation, team orientation and mutual support. Team effectiveness is represented by EMCC operator's evaluation of performance satisfaction, situational awareness and mission complexity. The predicted directions of the associations are marked on the arrows (*Negative association between team processes and measure of team effectiveness).

item: "To what extent did the EMCC-operator and the ambulance personnel maintain the best possible overview of the situation?"

Mission complexity. Mission complexity was measured with the item: "To what extent did you perceive the mission as complex?"

Closed-loop communication and all five team process variables were measured using frequency counts derived by study personnel from audio recordings of the interactions between EMCC operators and ambulance personnel.

Closed Loop Communication. Scores on CLC were obtained from verbal statements of the logged communication between the EMCC-operator and the ambulance personnel during the execution of the real-life emergencies. The number of confirmations acted as a measure of the third coordinating mechanism.

Team processes. By analyzing the logged authentic communication between the EMCC-operator and the ambulance personnel, frequencies of the "Big five" team behavior described by Salas et al. (2005) were recorded. **Table 1** shows the five team processes including a generic description and examples. The frequencies were summed making an aggregated score for team processes.

Procedure

According to Norwegian regulations, ethical approval for anonymous non-health-related data or quality improvement studies is not required. However, a formal letter of acceptance confirmed the approval of the study by the EMCC-administration, based on regulations in the Health Personnel Act (Helsepersonelloven) §26. Since a prerequisite for conducting the study was that data collection should not intervene with ongoing activity at the EMCC, several meetings were held with key personnel to develop the short questionnaire and entertain mutual understanding. Verbal and written information was presented to the EMCC-operators with a focus on anonymity and consent. Only written information was presented to the ambulance personnel.

As a standard procedure, all communication, internal in the EMCC and external toward the ambulances, are stored in a voice-recorded database. Thus, all voice recordings included in the present study were based on real-life ambulance missions. The authentic voice recordings consist of complete audio recordings from the emergency call to the end of the mission. Since the study was based on real life emergencies, team composition (e.g., EMCC operator and ambulance workers) were done using standard protocol for handling emergency calls in the EMCC. This includes random assignment to team based on availability of EMCC-operator and ambulance, distance to the emergency and urgency of need for care.

By retrieving the logged verbal communication, the behavioral markers of team processes and CLC could be scored. Based on the unique identification number of the mission, the anonymous data from the questionnaires were combined with an anonymized version of the stored communication (i.e., the names of the EMCC and ambulance personnel, and patient, were deleted). The variables were extracted by manually scoring pre-defined categories of verbal statements (see **Table 1**) from digital sound-files played on a computer. The scoring was done by three research assistants, trained in this specific methodology.

The questionnaire administered to the EMCC-operators was filled in immediately after completion of a mission.

Analyses

Reliability and correlational analyses were conducted using SPSS 25. Two-tailed Pearson product-moment correlation coefficients

TABLE 1 | Big-five team processes based on Salas et al. (2005) including behavioral markers, generic observed behavior and examples.

Team processes	Behavioral markers	Observed behavior	Examples
Team leadership	Coordination of team	Organizing, prioritizing and updating	Will arrive (at the E.R) in . . . minutes.
Team orientation	Team goals above own goals	Suggestions, recommendations, involvement in tasks	Patient reporting pain in . . ..region. Checking response and suggest . . ..
Monitoring	Assess performance	Controlling, checking and self-correction	Do you copy?» and «understood»?
Adaptation	Adjust behavior to environment	Change in plans	It might be best if you. . ..
Support	Predicting needs of team-members and share workload	Taking over tasks, offering resources and information	"Would you like to speak to the local GP?" I will make the necessary arrangements

were used in the bi-variate correlational analyses, and intra-class correlations were used to test the inter-rater reliability of the scoring of the coordinating mechanisms and the "Big five" team processes. The intra-class correlations were based on 25 voice recordings scored by all raters. Path analyses were conducted using AMOS 25, testing the *a priori* hypothesis model show in **Figure 1**. Separate analyses were performed for each of the three dependent variables of performance satisfaction, mission complexity and situational awareness. In the proposed model SMM, trust, and CLC acted as exogenous variables, while team processes and dependent variables acted as endogenous variables. The effects are based on generalized least square estimates. Ninety percent confidence intervals (90% CI) were determined using 1,000 bootstrapped samples. The proportion of explained variance in variables was calculated using multiple squared correlations. To compare the fit of the proposed model with the observed data, Comparative Fit Indexes (CFI) were computed for each dependent variable. A value close to 0.95 indicates a good fit between the hypothesized model and the observed data, whereas values in the range of 0.90–0.95 are considered acceptable (Hu and Bentler, 1999).

RESULTS

Descriptive Data

In the present study, 92% of the missions were coordinated by EMCC-operators with a background as ambulance paramedics, whereas 6% of the missions were coordinated by operators with a dual background as a registered nurse and ambulance paramedic (2% missing data).

The measure of "Big five" team processes showed an alpha-value of 0.593. This low alpha value indicated the multi-dimensionality of the measure. This is to be expected since the variable is a composite measure of five different team processes. Three raters scored the communications showing an intra-class correlation ranging from $r = 0.726$ to $r = 0.967$ on the different dependent measures. Means and standard deviations are presented in **Table 2**.

Relationship Between the Shared Mental Model Approach and Measures of Effectiveness

Positive correlations were found between SMM and the dependent measures of performance satisfaction and situational awareness (see **Table 2**). A significant negative correlation was revealed between SMM and complexity of the situation. The same pattern was found for the mechanism of trust. The only significant correlation involving CLC was a positive association to team processes, and the only inter-correlation between the mechanism was a relation between trust and SMM. These results partially support hypothesis 5.

The Fitness of the Shared Mental Model Approach in Explaining Team Effectiveness

Shared Mental Models and the Prediction of Performance Satisfaction

A significant positive direct effect from SMM to Performance satisfaction ($p < 0.001$; see **Table 3**) was found. Thus, supporting our first hypothesis. No direct effect of CLS on performance satisfaction was found. Thus, hypothesis 2 was not supported. The present analysis also revealed a positive total effect (direct and indirect) on the path from SMM to performance satisfaction, using team processes as a mediator ($p < 0.002$; 90% bootstrapped CI = 0.375 to 0.604). No paths flowing from CLC through team processes were found (not supporting hypothesis 4). The path analysis further showed significant positive paths from SMM and CLC to team processes ($p < 0.04$ and $p < 0.001$, respectively). No effect of team processes on performance satisfaction was revealed. The multiple squared correlation analyses showed that

TABLE 2 | Means (M), standard deviations and inter-correlations for all variables in the proposed model.

	M	SD	1	2	3	4	5	6
Shared Mental Model (1)	78.69	13.86						
Trust (2)	80.94	13.75	0.52**					
Closed Loop Comms. (3)	2.03	1.75	−0.06	−0.10				
Team processes (4)	7.06	7.27	0.04	−0.13	0.79**			
Performance satisfaction (5)	83.06	12.02	0.60**	0.78**	−0.13	−0.09		
Situational awareness (6)	77.24	15.04	0.56**	0.38**	0.08	0.18	0.52**	
Mission complexity	27.71	22.34	−0.46**	−0.43**	0.13	0.11	−0.33**	−0.09

***p < .01.*

TABLE 3 | Regression weights for the proposed paths in the model predicting the three dependent variables (DV) of subjective evaluation of performance satisfaction, situational awareness, and mission complexity (complexity).

Paths	Performance satisfaction		Situational awareness		Complexity	
	Un-standardized	β-weights	Un-standardized	β-weights	Un-standardized	β-weights
Direct effects						
SMM - > DV	0.523***	0.78	0.599***	0.561	−0.744***	−0.483
SMM - > Team processes	0.146*	0.278	0.087*	0.167	0.09*	0.171
Team Processes - > DV	−0.167	−0.131	0.375	0.185	0.381	0.13
Trust - > Team processes	−0.191	−0.269	−0.077	−0.143	−0.082	−0.148
Closed Loop Comms. - > DV	−0.118	−0.022	−0.261	−0.031	0.049	0.004
Closed Loop Comms - > Team processes	3.186***	0.767	3.247***	0.782	3.244***	0.781
Indirect effects						
SMM - > Team processes - > DV	−0.024	−0.036	0.033	0.031	0.034	0.022
Trust - > team processes - > DV	0.032	0.035	−0.029	−0.026	−0.031	−0.019
Closed Loop Comms. - > Team processes - > DV	−0.532	−0.1	1.219	0.144	1.236	0.101
Total effects						
SMM - > DV	0.498**	0.743	0.632**	0.592	−0.710**	−0.460
Closed Loop Comms. - > DV	−0.649	−0.122	0.958	0.113	1.286	0.105

Table presents direct, indirect and total effects. The effects are separated for unstandardized and standardized (β) weights.
**p < 0.05; **p < 0.01; ***p < 0.001.*

team mechanisms explained 64% ($R^2 = 0.64$) of the variance in team processes. Sixty-two percent of the variance in performance satisfaction was explained ($R^2 = 0.62$). When fitting the proposed model onto the observed data, a CFI of 0.573 occurred. This indicates a low fit for the model.

Shared Mental Models and the Prediction of Situational Awareness

The analysis using situation awareness as an outcome variable revealed a significant positive path from SMM directly to situational awareness ($p < 0.001$; see **Table 3**), supporting hypotheses 1. No effect involving CLC on situational awareness reached the level of significance. Thus, hypothesis 2 was not supported. The total effect of the path between SMM and situational awareness using team processes as mediator, was significant ($p = 0.02$, 90% bootstrapped confidence interval bounds between 0.456 and 0.800). No significant paths flowing from CLC through team processes on situational awareness was found.

The multiple squared correlation analyses revealed that 35% ($R^2 = 0.35$) of the variance in situational awareness was explained. Fitting the proposed model on the observed data obtained a CFI of 0.979, indicating a good fit of the model.

Shared Mental Models and the Prediction of Mission Complexity

SMM showed a negative direct effect on evaluation of mission complexity ($p < 0.001$; see **Table 3**), supporting hypothesis 1. No direct effect involving CLC on mission complexity reached the level of significance. Thus, hypothesis 2 was not supported. A total effect of the relation of SMM to complexity, with team processes as a mediator ($p < 0.002$; 90% bootstrapped confidence intervals between −0.951 and −0.457) was also found. No path flowing from CLC through team processes to mission complexity was revealed, providing no support to the fourth hypothesis. No effect of team processes on mission complexity was found. Furthermore, no effects of trust related to any of the measures of team effectiveness were found. The multiple squared correlation for complexity was 0.25, explaining 25% of the variance in mission complexity. An acceptable fit between the proposed model and the data was obtained (CFI = 0.928).

DISCUSSION

The path analyses showed a direct positive effect of SMM on the EMCC-operator's evaluation of the team's execution of the missions (performance satisfaction), the team's situational awareness, as well as a negative direct effect on perceived mission complexity. Furthermore, both SMM and CLC showed positive predictions of team processes measured as the sum of the "Big five" key team processes. The fit index was shown to vary from good to acceptable for predicting situational awareness and Mission complexity, while the fit for the performance satisfaction model was low.

Coordinating Mechanisms and Emergency Medical Teams Effectiveness

It is argued that virtual teams (i.e., dispersed team members) operate more autonomously, requiring a higher quality of intra-team teamwork (Johnson et al., 2011). The SMM approach suggests a conceptual model of teamwork characterized by a causal flow from coordinating mechanisms through team processes to performance (Salas et al., 2005). For instance, Mathieu et al. (2000) found that SMM was positively related to team processes and, subsequently through a complete indirect effect, performance in co-located flight simulator teams. However, the prediction of team processes as mediators between

mechanisms and performance was not supported in the present study. The results from the path analysis revealed that the EMS-teams' SMM enhances the evaluation of performance satisfaction and situational awareness, while reducing perceived mission complexity. However, this occurred without the mediating role of the "Big five" team processes. One reason for the findings could be that the transition from face-to-face communication to teamwork in virtual teams actually alters the underlying mechanism, causing successful execution of the team's mission by increasing the importance of shared cognitions representing vital aspects of the mission. The "Big five" approach describes various processes with behavioral markers that are important for the face-to-face interaction among a co-located team. For a physically dispersed team, there is a need for additional processes to coordinate the team and to establish a shared understanding of task, team and occurring challenges. When team members are separated, the team competencies are reduced, relying more on mechanisms that aims at distributing information equally among team members in order to create sharedness within the team. This need for shared understanding could be fueled by a requirement within the team in order to adapt their behavior to higher order systems. Pre-hospital missions often involve standard medical procedures and a need for integrating these procedures in the medical evacuation chain (i.e., higher order systems). This creates requirements of a shared cognitive representation of the status at the site of the emergency as well as projection in the near future and the requirements are met by the coordinating mechanism of shared mental model. Thus, the need for a shared understanding in order to integrate and coordinate the team's activity with present and future intentions of the team, guides team interaction based on SMM while verbal markers of team processes are reduced. Our claim is that the increased need for coordinating mechanisms, and especially SMM, could be generalized to other virtual teams. Geographically dispersed team members would perform their professional procedures, but in order to coordinate their activities at present and toward future states, mechanisms that creates sharedness of understanding is crucial. In the absence of visual cues, verbal communication becomes crucial to establish shared mental models.

The main finding of shared mental model being the only predictor of performance lends some support also from studies using a related research approach. Shared mental model is characterized by a shared understanding in the team of tasks, procedures, team members and equipment (Jonker et al., 2011; Sinval et al., 2020). The shared understanding approach shows similar characteristics and refers to the degree to which people concur on the interpretation of significant concepts, when sharing a perspective (mutual agreement) or can act in a coordinated manner (Agredo-Delgado et al., 2020). Agredo-Delgado et al. (2020) investigated the development of shared understanding in teams using computer-supported collaborative work. They found a marked effect on both sharedness within the group as well as performance toward the objectives when focusing on three phases of teamwork. The pre-process phase was composed of activity related to design and specification of objectives and possible disagreements followed by the execution of collaborative activity (process phase). In the post-process phase, a review of the process related to specific objectives was conducted. The design of the Agredo-Delgado et al. (2020) study indicated a mediating role of team processes, Thus, an abundance of empirical data from different research approaches supports the importance of shared cognitive representation of objectives, procedures and team interaction. However, a discrepancy exists of the mediating role of team processes. One reason for the discrepancy could be found in the type of tasks studied. The present study differs from most of the research using the "Big five" approach by studying teams conducting real life critical missions.

Significant aspects when evaluating team performance is the team members' subjective experiences of mission success. Important factors when rating mission success are the team's ability to fulfill their objectives when all aspects are considered (i.e., performance satisfaction), gain a correct understanding of the situation (i.e., situational awareness), and obtain clarity regarding the situation (i.e., complexity). SMM being positively related to the evaluation of team situational awareness is also in line with previous research (Macintosh et al., 2009; Saus et al., 2010; Johnsen et al., 2017). Macintosh et al. (2009) found team SA to be related to decision making in face-to-face teamwork in United Kingdom delivery suits. In the present study, this was shown to also be the case for team SA in dispersed medical teams. In a review of teamwork in primary care, Fiscella and McDaniel (2018) emphasized enabling factors for frequent face-to face communication (e.g., space configurations) in order to generate and maintain shared mental models, shared goals, and shared decision making in primary care teams. The negative association between SMM and complexity shows that increased levels of SMM results in a decreased evaluation of the complexity of the situation. This is a notable finding since a perception of reduced situation complexity (e.g., simplicity) may increase teams' recognition-based decisions (Drury et al., 2012) and prompt the use of standard operating procedures. In a study in the medical domain using the SMM approach, Schraagen (2011) reported complexity as a major risk factor in non-routine pediatric cases. Complexity was related to longer operations as well as more negative patient outcome. However, the study reported that one of the teams studied encountered a more severe case a few weeks after the first, resulting in improved teamwork, higher degrees of attending and employing procedures in a heedful way (Schraagen, 2011). Taken together, the present study expands previous knowledge by showing that SMM is the only team behavior predicting EMS team effectiveness in virtual teams conducting real life missions. Coordinating mechanisms and "The Big-five" team processes.

The results show that the mechanisms of SMM and CLC predict the "Big-five" team processes. This is in line with the predictions of previous models (Salas et al., 2005), where the purpose of the coordinating mechanisms was to support an even distribution of information, fueling the "Big five" team processes resulting in enhanced performance (Uitdewilligen et al., 2018). The effect of team processes on team effectiveness has previously been reported in the medical domain. Studies of team processes within the nursing role (Kalisch et al., 2009), as well as studies of Norwegian trauma teams (Westli et al., 2010; Johnsen et al., 2017), have reported the same predicted effect. However, a common factor in these studies was the co-location of team members involving face-to-face contact. This influence of team

process on performance was not supported in our study involving dispersed EMS teams. The direct effect of SMM outperforming an effect mediated by team processes could be caused by an increased need for shared knowledge structures when the team members are located separately.

The present study showed a high correlation between trust and SMM, which is in line with previous findings (Hanna and Richards, 2018). While both SMM and CLC were related to team processes, trust was not. McComb et al. (2017) compared nurses' and physicians' level of SMM and trust. They reported a low level of SMM regarding the perceived role responsibilities between nurses and physicians. The two professions showed an equal level of trust toward physicians, while the physicians rated their trust toward nurses lower compared nurses' evaluation of trust toward their own profession. A seemingly surprising discovery in the present study was the finding of no relation between trust and team processes. Trust could be defined as a belief that the team members will perform expected actions and recognize and protect the interests of their colleagues, as well as a willingness to allow for risk-exposure among members working interdependently (Salas et al., 2005). A possible explanation could be that both CLC and SMM are mechanisms directly involved in generating, maintaining, and altering knowledge structures in the teams, while the impact of trust flows through an indirect emotional component resulting in team members acceptance of being monitored. In addition, for teams consisting of EMCC-operators and paramedics, the function of SMM and CLC are more easily communicated verbally, while signals of trust or lack of trust are harder to communicate and perceive when not being in visual contact. Thus, separating team members geographically alters the mechanisms underlying teamwork and team effectiveness by extinguishing the role of trust. The model fit of the tested model varied from poor (performance satisfaction), to acceptable (mission complexity), to good (situational awareness). This indicates that the model needs to be improved to show more consistency over different indices of team effectiveness.

Implications

The result from the present study shows the importance of the coordinating mechanism of SMM. Four types of SMMs have been proposed (Cannon-Bowers et al., 1995; Cannon-Bowers and Salas, 1998). This includes shared cognition of equipment, operational task, interaction as well as of the team members. Based on the SMM approach, cross-training is suggested to present an important method to increase shared team knowledge structures (McCann et al., 2000; Marks et al., 2002). Cross training refers to a rotation of team roles following training in one's own role requirements. This provides an opportunity to experience new learning for all team members and facilitate information sharing that will extend and maintain SMMs. Superior training effects have been found in teams characterized by high levels of SMMs. Espevik et al. (2011b) reported significantly higher levels of performance after only one training session in naval cadet teams exposed to cross-training. Implications of the present study point not only toward the type of training strategy (e.g., cross training), but also toward the training objectives (types of SMM). Thus, in targeting future training it will be paramount to define the type of SMM in need of training ("What") and the training strategy ("How") to increase shared cognitions in EMS teams.

Limitations

Some limitations should be noted. The present study relies on single item measures on several of the variables. From a historical standpoint, reliability and validity issues have discouraged the use of single item measures of psychological constructs. Thus, multiple-item scales are favored (Nunnaly, 1967). However, this view has been challenged and there are several reports supporting the use of single-item variables (Wanous and Reichers, 1996; Wanous et al., 1997). This is founded on empirical data showing high test -retest reliability (Gardner et al., 1998), as well as well as high correlations with multiple item scales (Littman et al., 2006). The validity is also revealed by single item measures effectively predicting outcomes (Littman et al., 2006), including job satisfaction (Wanous et al., 1997; Nagy, 2002). Although there are limitations to single-item measures, potential advantages should also be noted. Advantages like cost-efficiency, greater face validity, and a possible increased willingness of respondents to take time to complete the questionnaire instigated by a less intrusive method compared to the use of multi-item scales.

The item measuring situation awareness is not totally in line with Endsley's model of situational awareness (Endsley, 1995). The model describes situational awareness as being constituted by three hierarchical organized levels. These levels include detection of critical signals (level 1), understanding of the situation (level 2) and prediction into the near future (level 3). However, the question is meant to capture the core element of the concept by tapping into the individual's cognitive representation and understanding of the emergency at hand. It could be argued that that this includes level 1 and level 2 in Endsley's (1995) model.

Both the dependent variables and two of the coordinating mechanisms were evaluated by the EMCC-operators only. This was caused by the study's intention to not intrude on the ongoing activities. Logistical, methodological, and ethical constraints prevented recordings of data from the ambulance personnel or regarding outcome for the patients. However, the EMCC-operator is a significant team member. By acting as a communication hub, initiating and coordinating the mission, the operators could be viewed as team leaders, and leaders' evaluation of team effectiveness is imperative.

Another limitation is the lack of control concerning the team members' identification with the unit studied. The results could be influenced by the participants defining their team membership differently. The ambulance crews could view themselves as part of one team and the EMCC-operators as another team, resulting in measuring inter-team coordination rather than teamwork. Despite this, the units that were studied match the formal definition of teams, justifying the approach of the current study. Finally, nations vary in the organization of their pre-hospital services. Education and training levels, the size of organizations, and the amount and type of operations may differ, limiting the generalizability of the results.

CONCLUSION

A consequence of geographically dispersed EMS teams is a change of effective team behavior compared to face-to-face teamwork. The present study reveals a lack of influence with respect to team processes on effectiveness when EMS team members are dispersed. This leads to an increased emphasis on the coordinating mechanism of SMM, demonstrating a direct effect on all the dependent variables of performance satisfaction, situational awareness and mission complexity. The lack of effects on team processes could be due to all team processes being channeled *via* verbal communications only, as well as a need for shared understanding in order to integrated and coordinate the team's behavior in a higher order system. In this case the evacuation chain from emergency site to higher echelon care. This leaves virtual EMS teams with a high degree of dependency on coordinating mechanisms, for which SMM seem to be crucial. This is important knowledge when determining education and training aimed at increasing the collaboration between EMCC-operators and the ambulance personnel to improve patient safety. However, more research is needed to understand the unique roles of the different types of SMM in predicting performance in different types of teams.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

REFERENCES

Agredo-Delgado, V., Ruiz, P. H., Mon, A., Collazos, C. A., Moreira, F., and Fardoun, H. M. (2020). "Validating the shared understanding construction in computer supported collaborative work in a problem-solving activity," in *Trends and Innovations in Information Systems and Technologies*, eds A. Rocha, H. AdeliLuís, P. Reis, S. Costanzo, I. Orovic, and F. Moreira (Cham: Springer), 203–214. doi: 10.1007/978-3-030-45697-9_20

Algesheimer, R., Dholakia, U. M., and Gurau, C. (2011). Virtual team performance in a highly competitive environment. *Group Organ. Manag.* 36, 161–190.

Cannon-Bowers, J. A., and Salas, E. (1998). *Making Decisions Under Stress*. Washington, DC: APA.

Cannon-Bowers, J. A., Tannenbaum, S. I., Salas, E., and Volpe, C. E. (1995). "Defining competencies and establishing team training requirements," in *Team Effectiveness and Decision Making in Organizations*, eds R. A. Guzzo, E. Salas, and Associates (San Francisco, CA: Jossey-Bass), 333–380.

Curseu, P. L., Schalk, R., and Wessel, I. (2008). How do virtual teams process information? A literature review and implications for management. *J. Manag. Psychol.* 23, 628–652.

DeChurch, L. A., and Mesmer-Magnus, J. R. (2010). The cognitive underpinnings of effective teamwork: a meta-analysis. *J. Appl. Psychol.* 95, 32–53. doi: 10.1037/a0017328

Drury, M., Conboy, K., and Power, K. (2012). Obstacles to decision making in Agile software development teams. *J. Syst. Softw.* 85, 1239–1254. doi: 10.1016/j.jss.2012.01.058

El-Shafy, I. A., Delgado, J., Akerman, M., Bullaro, F., Christopherson, N. A., and Prince, J. M. (2018). Closed-loop communication improves task completion in pediatric trauma resuscitation. *J. Surg. Educ.* 75, 58–64. doi: 10.1016/j.jsurg.2017.06.025

Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. *Hum. Fact.* 37, 32–64. doi: 10.1518/001872095779049543

Espevik, R., Johnsen, B. H., and Eid, J. (2011a). Communication and performance in co-located and distributed teams: an issue of shared mental models of team members? *Milit. Psychol.* 23, 616–638. doi: 10.1080/08995605.2011.616792

Espevik, R., Johnsen, B. H., and Eid, J. (2011b). Outcomes of shared mental models of team members in cross training and high intensity situations. *J. Cogn. Eng. Dec. Mak.* 5, 352–377. doi: 10.1177/1555343411424695

Espevik, R., Johnsen, B. H., Eid, J., and Thayer, J. (2006). Shared mental models and operational effectiveness: effects on performance and team processes in a submarine attack team. *Milit. Psychol.* 18, 23–36.

Fiscella, K., and McDaniel, S. H. (2018). The complexity, diversity, and science of primary care teams. *Am. Psychol.* 73, 451–467. doi: 10.1037/amp0000244

Fletcher, G., Flin, R., McGeorge, P., Galvin, R., Maran, N., and Patey, R. (2003). Anesthetists' Non-Technical Skills (ANTS): evaluation of a behavioural marker system. *Br. J. Anaesth.* 90, 580–588.

Flin, R., and Maran, N. (2004). Identifying and training non-technical skills for teams in acute medicine. *Q. Saf. Health Care* 13, 80–94. doi: 10.1136/qshc.2004.009993

Gardner, D. G., Cummings, L. L., Dunham, R. B., and Pierce, J. L. (1998). Single-item versus multiple-item measurement scales: an empirical comparison. *Educ. Psychol. Meas.* 58, 898–915. doi: 10.1177/0013164498058006003

General Medical Council (2016). *Tomorrow's Doctors: Outcomes and Standards for Undergraduate Medical Education*. Available online at: www.gmc-uk.org/Tomorrow_s_Doctors_1214.pdf_48905759.pdf (accessed December 16, 2016).

Hanna, N., and Richards, D. (2018). The impact of multimodal communication on a shared mental model, trust, and commitment in human–intelligent virtual agent teams. *Multimod. Technol. Interact.* 2, 1–27. doi: 10.3390/mti2030048

Härgenstam, M., Lindkvist, M., Brulig, C., Jacobsson, M., and Hultin, M. (2013). Communication in interdisciplinary teams: exploring closed-loop communication during in situ trauma team training. *BMJ Open* 3:e003525. doi: 10.1136/bmjopen-2013003525

Hu, L., and Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct. Equ. Model.* 6, 1–55. doi: 10.1080/10705519909540118

Johnsen, B. H., Westli, H. K., Espevik, R., Wisborg, T., and Brattebø, G. (2017). High performance trauma teams: frequency of behavioral markers of shared mental model displayed by team leaders and quality of performance. *Scand. J. Trauma Resusc. Emerg. Med.* 25:109. doi: 10.1186/s13049-017-0452-3

Johnson, T. E., Top, E., and Yukselturk, E. (2011). Team shared mental model as a contributing factor to team performance and students' course satisfaction in blended courses. *Comput. Hum. Behav.* 27, 2330–2338. doi: 10.1016/j.chb.2011.07.012

Johnson-Laird, P. (1983). *Mental Models*. Cambridge, MA: Harvard University Press.

Jonker C. M., van Riemsdijk, M. B., and Vermeulen, B. (2011). "Shared mental models," in *Coordination, Organizations, Institutions, and Norms in Agent Systems VI. COIN 2010. Lecture Notes in Computer Science*, vol 6541, eds M.

De Vos, N. Fornara, J. V. Pitt, G. Vouros. Springer, Berlin, Heidelberg. doi: 10.1007/978-3-642-21268-0_8

Jouanne, E., Charron, C., Chauvin, C., and Morel, G. (2017). Correlates of team effectiveness: an exploratory study of firefighter's operations during emergency situations. *Appl. Ergon.* 61, 69–77. doi: 10.1016/j.apergo.2017.01.005

Kalisch, B. J., Weaver, S. J., and Salas, E. (2009). What does nursing teamwork look like? a qualitative study. *J. Nurs. Care Q.* 24, 298–307. doi: 10.1097/NCQ.0b013e3181a001c0

Kraiger, K., and Wenzel, L. H. (1997). "Conceptual development and empirical evaluation of measures of shared mental models as indicators of team effectiveness," in *Team Performance Assessment and Measurement: Theory, Methods and Applications*, eds M. T. Brannick, E. Salas, and C. Prince (Mahwah, NJ: Lawrence Erlbaum), 63–84.

Littman, A. J., White, E., Satia, J. A., Bowen, D. J., and Kristal, A. R. (2006). Reliability and validity of two single-item measures of psychosocial stress. *Epidemiology* 17, 398–403. doi: 10.1097/01.ede.0000219721.89552.51

Macintosh, N., Berridge, E. J., and Freeth, D. (2009). Supporting structures for team situational awareness and decision making: insights from four delivery suits. *J. Eval. Clin. Pract.* 15, 46–54. doi: 10.1111/j.1365-2753.2008.00953.x

Marks, M. A., Sabella, M. J., Burke, C. S., and Zaccaro, S. J. (2002). The impact of cross-training on team effectiveness. *J. Appl. Psychol.* 87, 3–13. doi: 10.1037/0021-9010.87.1.3

Mathieu, J. E., Heffner, T. S., Goodwin, G. F., Salas, E., and Cannon-Bowers, J. A. (2000). The influence of shared mental models on team process and performance. *J. Appl. Psychol.* 85, 273–283. doi: 10.1037/0021-9010.85.2.273

McCann, C., Baranski, J. V., Thompson, M. M., and Pigeau, R. A. (2000). On the utility of experimental cross-training for decision making under time stress. *Ergonomics* 43, 1095–1110.

McComb, S. A., Lemaster, M., Henneman, E. A., and Hinchey, K. T. (2017). An evaluation of shared mental models and mutual trust on general medical units: implications for collaboration teamwork and patient safety. *J. Patient Saf.* 13, 237–242. doi: 10.1097/PTS.0000000000000151

McCormack, H. M., David, J. D. L., and Sheather, S. (1988). Clinical applications of visual analogue scales: a critical review. *Psychol. Med.* 18, 1007–1019. doi: 10.1017/s0033291700009934

Mesmer-Magnus, J., Niler, A. A., Plummer, G., Larson, L. E., and DeChurch, L. A. (2017). The cognitive underpinnings of effective teamwork: a continuation. *Career Dev. Intern.* 22, 507–519. doi: 10.1108/cdi-08-2017-0140

Meynard, M. T., and Gilson, L. L. (2014). The role of shared mental model development in understanding virtual team effectiveness. *Group Organ. Manag.* 39, 3–32.

Mjelde, F. V., Smith, K., Lunde, P., and Espevik, R. (2015). Military teams – a demand for resilience. *Work* 54, 283–294. doi: 10.3233/WOR-162298

Nagy, M. (2002). Using a single-item approach to measure facet job satisfaction. *J. Occup. Organ. Psychol.* 75, 77–86. doi: 10.1348/096317902167658

Nicholls, A. R., Polman, R. C. J., and Levy, A. R. (2012). A path-analysis of stress appraisals. Emotions, coping, and performance satisfaction among athletes. *Psychol. Sport Exerc.* 13:263e270. doi: 10.1016/j.psychsport.2011.12.003

Norwegian Medical Association (2009). *Norsk Indeks for Medisinsk Nødhjelp (Norwegian Index for Medical Emergency Assistance)*, 3rd Edn, Stavanger: Laerdal Medical A/S.

Nunnaly, J. C. (1967). *Psychometric Theory*. New York, NY: McGraw-Hill.

Parelta, C. F., Lourenco, P. R., Lopes, P. N., Baptista, C., and Pais, P. N. (2018). Team development: definition, measurement and relationships with team effectiveness. *Hum. Perform.* 31, 97–124. doi: 10.1080/08959285.2018.1455685

Rentsch, J. R., and Klimoski, R. (2001). Why do "great minds" think alike? Antecedents of team member schema agreement. *J. Organ. Behav.* 22, 107–120. doi: 10.1002/job.81

Salas, E., Sims, D. E., and Burke, C. S. (2005). Is there a "Big Five" in teamwork. *Small Group Res.* 36, 555–599. doi: 10.1177/1046496405277134

Saus, E. R., Espevik, R., and Eid, J. (2010). "Situational awareness and shared mental models: implications for training in security operations," in *Enhancing Human Performance in Security Operations: International and Law Enforcement Perspectives*, eds P. T. Bartone, B. H. Johnsen, J. Eid, J. M. Violanti, and J. C. Laberg (Springfield, IL: Charles C. Thomas, Publishers), 143–161.

Schraagen, J. M. (2011). Dealing with unforeseen complexity in the OR: the role of heedful interrelating in medical teams. *Theoret. Issues Ergon. Sci.* 12, 256–272. doi: 10.1080/1464536X.2011.564481

Sinval, J., João, M., Marques, S. C., Uitdewilligen, S., Travis, M. M., and Margarida, P. A. (2020). Development of the Referee Shared Mental Models Measure (RSMMM). *Front. Psychol.* 11:550271. doi: 10.3389/fpsyg.2020.550271

Uitdewilligen, S., Rico, R., and Waller, M. J. (2018). Fluid and stable: dynamics of team action patterns and adaptive outcomes. *J. Organ. Behav.* 39, 1113–1128. doi: 10.1002/job.2267

UK National Health Service (2010). *Institute for Innovation and Improvement. Medical Leadership Competency Framework; Enhancing Engagement in Medical Leadership, July; Third Edition*. Available online at: https://www.leadershipacademy.nhs.uk/wp-content/uploads/2012/11/NHSLeadership-Leadership-Framework-Medical-Leadership-Competency-Framework-3rd-ed.pdf (accessed December 16, 2016).

Undre, S., Sevdalis, N., Healey, A. N., Darzi, S., and Vincent, C. A. (2006). Teamwork in the operating theatre: cohesion or confusion?. *J. Eval. Clin. Pract.* 12, 182–189. doi: 10.1111/j.1365-2753.2006.00614.x

Wanous, J. P., and Reichers, A. E. (1996). Estimating the reliability of a single-item measure. *Psychol. Rep.* 78, 631–634.

Wanous, J. P., Reichers, A. E., and Hudy, M. J. (1997). Overall job satisfaction: how good are single-item measures?. *J. Appl. Psychol.* 82, 247–252. doi: 10.1037/0021-9010.82.2.247

Weld, L. R., Stringer, M. T., Ebertowski, J. S., Baumgartner, T. S., Kasprenski, M. C., Kelley, J. C., et al. (2016). TeamSTEPPS improves operating room efficiency and patient safety. *Am. J. Med. Q.* 31, 408–414. doi: 10.1177/1062860615583671

Westli, H. K., Johnsen, B. H., Eid, J., Rasten, I., and Brattebø, G. (2010). Teamwork skills shared mental models, and performance in simulated trauma teams. *Scand. J. Trauma Resusc. Emerg. Med.* 18, 18–47. doi: 10.1186/1757-7241-18-47

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Johnsen, Espevik, Eid, Østerås, Jacobsen and Brattebø. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Psychology

ORIGINAL RESEARCH
published: 24 March 2022
doi: 10.3389/fpsyg.2022.857647

Prosodic Phrasing of Good Speakers in English and Czech

Radek Skarnitzl[1*†] *and Hana Hledíková*[2†]

[1]*Institute of Phonetics, Faculty of Arts, Charles University, Prague, Czechia,* [2]*Department of English Language and ELT Methodology, Faculty of Arts, Charles University, Prague, Czechia*

Prosodic patterning is known to affect the impression that speakers make on their listeners. This study explores prosodic phrasing in good public speakers of American English and Czech. Czech is a West Slavic language whose intonation is reported to be flatter and prosodic phrases longer than in English. We analyzed prosodic characteristics of 10 speakers of Czech and American English who appeared in TED Talks, assuming such appearance to be a mark of a "good speaker." Our objective was to see whether prosodic phrasing will be more similar in these public speeches between the two languages. We measured the length of prosodic phrases, speech rate in each phrase, and pitch range and melodic variability in the entire phrase, as well as in its pre-nuclear and nuclear portion. The number of syllables per phrase was higher in Czech than in English, although phrases were generally very short in both languages. The melodic indicators confirm smaller melodic variability in Czech even in publicly performed TED Talks. Overall, our results show that there are differences between Czech and English prosodic phrasing in good public speakers but that the genre also affects phrasing. Prosodic rendition—especially prosodic phrase length and melodic variability—is therefore a vital, albeit somewhat language-specific aspect of speech performance which public speakers should pay close attention to, both in their native language and in foreign languages.

Keywords: prosodic phrasing, pitch range, melodic variability, speaking rate, Czech, English

OPEN ACCESS

Edited by:
Oliver Niebuhr, University of Southern Denmark, Denmark

Reviewed by:
Plinio Almeida Barbosa, State University of Campinas, Brazil
Isabella Poggi, Roma Tre University, Italy

****Correspondence:***
Radek Skarnitzl
radek.skarnitzl@ff.cuni.cz

†These authors have contributed equally to this work and share first authorship

Specialty section:
This article was submitted to Language Sciences, a section of the journal Frontiers in Psychology

Received: *18 January 2022*
Accepted: *01 March 2022*
Published: *24 March 2022*

Citation:
Skarnitzl R and Hledíková H (2022) Prosodic Phrasing of Good Speakers in English and Czech. Front. Psychol. 13:857647. doi: 10.3389/fpsyg.2022.857647

INTRODUCTION

The study of personal charisma has a long history. Originally, charisma was regarded as a supernatural quality, an in-born talent granted only to few (Weber, 1947). While this idea was soon abandoned and we know today that charisma is a skill which can be acquired and trained (e.g., Towler, 2003; Antonakis et al., 2011; Niebuhr et al., 2019; Niebuhr and Neitsch, 2020), a number of myths surrounding the nature of charisma persist (Michalsky and Niebuhr, 2019). It has been established that charisma has a positive impact on, among others, the electoral success of politicians (Sheafer, 2008), the willingness of external stakeholders to participate in a company (Fanelli and Misangyi, 2006), university students' motivation and perception of their learning (Bolkan and Goodboy, 2014), the credibility of an advertised product (Gélinas-Chebat et al., 1996) or, sadly but not surprisingly, the attractiveness of radical preachers for their followers (Gendron, 2017).

It is an undeniable fact that charisma is closely related to communication skills. Several studies have compared the relative effect of content and delivery on the perception of charisma. In one of the first experimental studies, Holladay and Coombs (1994) combined visionary vs. non-visionary content with strong vs. weak delivery and found that strong delivery outweighed the effect of non-visionary content; in other words, delivery contributed more to the perception of charisma than content. Similar results were reported by Awamleh and Gardner (1999), who added organizational performance as a third factor determining a business leader's charisma. Amon (2016) cited in Michalsky and Niebuhr (2019, p. 36) goes so far as to claim that "[t]he moment you open your mouth, all the visible elements become mere decoration." In a more recent study concerning content and delivery, Caspi et al. (2019) extended the findings by offering a dual-process perspective: delivery is a stronger determiner of charisma because it is processed faster, more automatically than content whose processing requires conscious deliberation. According to the authors, the immediate impression based on the speaker's delivery "anchors" the perception of the speaker's charisma, which may only be modified if delivery and content are markedly misaligned.

In the above-mentioned studies, the components of delivery—apart from factors related directly to speech—included the maintenance of eye contact, gesticulation, and facial expressiveness, as well as relaxation and confidence of the speaker. While all these are key determinants of charismatic delivery, we will continue to focus only on speech characteristics. It is worth pointing out here, however, that the concept of charismatic speaker and charismatic speech is, to a certain extent, culture- or language-specific (Biadsy et al., 2008; D'Errico et al., 2013), and reported results should not be regarded as applicable universally.

For a long time, the advice featured in various manuals on rhetoric relied on impressionistic terms rather than on phonetic research; as shown by Niebuhr et al. (2017), however, respondents imagine very diverse concepts when describing the ideal voice as "rich," "durable," or "flexible." It is only with advancing phonetic research of the last approximately two decades that the label "charismatic voice" has been acquiring more specific meaning, and rhetoric manuals have been able to provide more explicit and targeted advice (see Gilner, 2014; Lucas, 2015: Ch. 13; Lower Mekong Initiative, 2017: Unit 7).

Charismatic speech has been researched in relation to both segmental and prosodic aspects, with a rather strong bias for the latter. In a segmentally oriented study, Niebuhr (2017) found that a moderate degree of reductions in speech resulted in speakers being perceived as more sociable, composed and sincere (traits which correlate with speaker charisma) than in speech which featured no reductions, as well as strong reductions. Interestingly, this result thus lends only limited support to the popular adage "Speak clearly!" On the other hand, two studies comparing Steve Jobs (famous for his strong charisma) and Mark Zuckerberg (known rather for the opposite) found clearer articulation of stop consonants (Niebuhr et al., 2018b) and vowels (Niebuhr and Gonzalez, 2019) in the late Apple CEO, while Zuckerberg's speech was characterized by less clear realizations of consonants and by a smaller vocalic space, respectively.

The above-mentioned exceptions notwithstanding, most research on charismatic speech has focused on prosodic characteristics or, as these are frequently called in rhetoric manuals, the tone of voice. Indeed, it is not surprising that good prosody has been recognized as contributing to high-quality speech the most; we may observe a parallel with second language pronunciation, where prosody has also become regarded as most beneficial in achieving learners' intelligibility (Derwing and Rossiter, 2003). As prosody in good speakers is the topic of the present study, we will examine the relation between charismatic speech and prosody in more detail in the next section.

Charismatic Speech and Prosody

Prosodic features of speech include in the perceptual domain, melodic organization, temporal characteristics (e.g., speech rate, rhythm, or the length of prosodic constituents), loudness, and voice quality, and their respective correlates in the acoustic domain, fundamental frequency (f_0), duration, intensity, and spectral characteristics. It is especially features of the first two groups—melodic and temporal aspects of speech—that have been shown to be crucial in the process of communication, playing an important role in the impression the speaker makes on listeners.

Most of the prosodic characteristics of a good speaker are intuitive; in other words, it is not surprising which kinds of prosodic behavior boost a speaker's charisma and which are detrimental to it. However, this does not apply to one's pitch level in a straightforward way—there seems to be no *a priori* reason why either a lower or higher pitch level should positively affect a speaker's charisma. Many rhetorical manuals advise their readers to lower the pitch of their voice: Barker (2011, p. 14) associates lower pitch with easier rapport creating and higher pitch with stress and tenseness in the voice; even the most current edition (Barker, 2019) advocates lower pitch repeatedly. Despite this, the majority of empirical research reports a positive correlation between speaking fundamental frequency and a speaker's charisma. Higher pitch level is therefore associated with higher and not lower charisma ratings. This was found in numerous studies which focused on charisma in politicians, for example, in Jacques Chirac's speeches (Touati, 1993), the debates of members of the Swedish parliament (Strangert and Gustafson, 2008), in the speech of nine candidates for the Democratic Party's 2004 presidential nomination (Rosenberg and Hirschberg, 2009), in a comparison of a French and Italian politician (D'Errico et al., 2013). A positive relationship between higher speaking fundamental frequency and perceived charisma was extended to business contexts: in a detailed analysis of Steve Jobs' voice, Niebuhr et al. (2016) found a markedly higher pitch level in Steve Jobs than in the reference speakers. It must be emphasized, however, that the higher pitch level does not consist in a mere upscaling of f_0 values: as found by Mixdorff et al. (2018, p. 817) in another comparison of Jobs and Zuckerberg, the higher f_0 mean may consist in the lowering of the f_0 baseline "and modifying the f_0 slopes of pitch accents and initial and final boundary tones such that they get longer, higher, and arrive faster at a high f_0 level" (*ibid*).

The latter finding brings us to f_0 range, a crucial factor in the creation of a speaker's charisma. According to all of the studies cited in the previous paragraph and others (e.g., Biadsy et al., 2008; Niebuhr et al., 2018a), higher charisma is associated with a larger melodic range. Given the nature of the speakers reported in this paper, it is worth mentioning a study by Berger et al. (2019), in which prosodic characteristics of the speech of YouTube video creators were correlated with the number of subscribers of their channels and the number of views and likes received. A significant positive correlation with f_0 range was only found with the first indicator, subscriber count. It is also important to draw attention to what Niebuhr et al. (2019) call overdose thresholds: there are limits to the positive correlations. One may imagine that an excessively large pitch range will make the speaker sound affected and untrustworthy; given the different habitual f_0 range in different languages (Mennen et al., 2012; Andreeva et al., 2014; Volín et al., 2015), the threshold is also likely to be language- or culture-specific (Grabe et al., 2003; Chen et al., 2004).

Another aspect which is related to melodic patterning and which has been found to correlate with perceived speaker charisma is focused words, or emphatic accents. Specifically, it appears that a higher number of these emphatic accents, as well as higher melodic peaks in them yield higher charisma ratings (Strangert and Gustafson, 2008). In addition, Niebuhr et al. (2016) documented not only a high number of instances of emphatic accentuation in Steve Jobs' speeches, but also all of its subtypes, with a prevalence of trust-seeking reinforcement and positive intensification.

Concerning the temporal organization of charismatic speech, faster speech rate appears to be preferred by listeners (Biadsy et al., 2008; Rosenberg and Hirschberg, 2009; Niebuhr et al., 2016), although it is obvious that an overdose threshold will apply in speech rate as well. In Steve Jobs' speech, Mixdorff et al. (2018) documented a strong and symmetric acceleration and deceleration in speech rate and a higher variability in speech rate, as compared with Mark Zuckerberg's speech.

Another key descriptor which is associated with good speakers is fluency, which may be defined as a natural flow of speech. According to De Jong and Mora (2019, p. 228), spoken fluency largely depends "on speakers' ability to execute the conceptualization and formulation of messages effectively and on their ability to translate formulated messages into articulatory actions smoothly during the speech production process." A number of fluency measures have been proposed, particularly in the domain of second language performance (Suzuki et al., 2021); some of these include articulation rate and speech rate, mean length of run between pauses, or the phonation time ratio, as well as the frequency of various disfluency phenomena such as hesitations, repetitions or repairs, prolongations, or false starts (McDougall and Duckworth, 2017). It is not surprising that more charismatic speakers produce fewer of such disfluencies, as shown for instance by Strangert and Gustafson (2008).

In their summary of good speaker characteristics, Niebuhr et al. (2017, p. 10) mention, with respect to fluency, that "inferred from the evidence we already have, the rhetorical advice to 'speak fluently' can be translated into 'split up your sentences into short phrases of no more than 4–5 words (about 2.0–2.5 s)'." Similarly, Niebuhr et al. (2019, p. 2) state in their summary of research that "prosodic-phrase durations… are negatively correlated with perceived speaker charisma." It is interesting, however, that there seems to be very little empirical evidence to support this claim. Some exceptions are the study by Strangert (2005), whose comparison of two skilled speakers revealed rather short phrases governed by the semantic rather than syntactic structure, as well as the comparison of Jobs and Zuckerberg by Niebuhr et al. (2016). Contrariwise, however, Rosenberg and Hirschberg (2009) report that the number of words per phrase was significantly and positively correlated with ratings of charisma.

It remains to be pointed out that there may be different types of charisma. For example, D'Errico et al. (2013) compared the speech of the American politicians Barack Obama and Donald Trump, as well as that of the Italian politicians Paolo Gentiloni and Matteo Salvini. While all of them may be regarded as charismatic in some sense, Obama's and Gentiloni's humility stands in stark contrast to Trump's and Salvini's dominance. The authors found marked differences between the humble and dominant politicians: the former used shorter utterances, more disfluencies, fewer stressed words, and speech which was less loud and slower than the latter.

To summarize the main findings of prosodic research into good speakers, "variation and variety are key concepts in creating a charismatic impact" (Niebuhr et al., 2016, p. 376). The validity of this adage is also supported by speech synthesis and resynthesis experiments: speech generated in a controlled way, using a large pitch range and few disfluencies, was rated more positively on charisma-related traits (Strangert and Gustafson, 2008; Fischer et al., 2019).

Prosodic Phrase as a Central Unit of Prosody

Prosodic features serve the role of organizing the flow of speech and giving it structure by dividing it into smaller units, which are called prosodic phrases, prosodic units, tone units, thought groups etc. The prosodic phrase is defined as "the domain of a perceptually coherent intonational contour" (Shattuck-Hufnagel and Turk, 1996, p. 210) and is delimited by prosodic boundaries. The strongest prominence of the prosodic phrase, realized on the tonic syllable, is called the nuclear stress. Nuclear stress tends to occur on the stressed syllable of the last content word in a phrase (Féry, 2017, p. 61) but may also be placed on other words to give it emphasis. Prosodic phrases are separated by prosodic boundaries, which are usually signaled by melodic and temporal features (melodic movements and final deceleration or lengthening, respectively), sometimes by a pause.

From the perspective of speech production, empirical evidence suggests that when speakers plan the lexical content of a stretch of speech, constructions which correspond in their size to a clause, with about four to six words, are preferred (Nooteboom, 1995; Pawley and Syder, 2000); it is not unreasonable to draw a parallel between a clause and prosodic phrase here. With

this one clause at a time hypothesis, Pawley and Syder point to the crucial role of a fluent unit which people can handle in a single focus of consciousness. In other words, prosodic phrasing tends to reflect the syntactic or information structure of a sentence (see also Cresti, 2018), but the authors claim that it is primarily not grammatical structure but a processing limit on our language planning that yields the typical prosodic structure of speech. We should point out that prosodic planning also depends on the capacity of speakers' working memory (Swets et al., 2007; Petrone et al., 2011).

From the perspective of speech perception, prosodic phrasing has been shown to play a significant role in listeners' comprehension. Numerous experiments which have been conducted since the 1960s and 70s showed that speech is processed faster and its contents are recalled with higher success rate if it is presented with clear phrasal prosody (e.g., O'Connell et al., 1968; Zurif and Mendelsohn, 1972; Leonard, 1974; Sturges and Martin, 1974; Reeves et al., 2000; Krivokapić, 2007). It therefore appears that it is phrasal prosody that provides the basic structure which allows us to hold a sequence of heard words in memory (Frazier et al., 2006).

Apart from being regulated by quasi-universal requirements of speech production and perception, prosodic phrasing differs to some extent between languages (Jun, 2003). Such differences may consist, for example, in the mapping between syntactic and prosodic structure, or in the effect of focus on the possibility of prosodic boundary placement. In this study, we analyze prosodic phrasing in "good speakers" of English and Czech; in the next section, we will therefore briefly compare the prosodic patterns of these two languages.

Prosodic Patterns in English and Czech

If we were to express the difference between English and Czech prosodic characteristics using a single impressionistic word, we could refer to English as vivid and to Czech as monotonous. At the level of individual words, lexical stress is manifested in English by higher f_0 level, longer duration, and shallower spectral slope (e.g., Eriksson and Heldner, 2015), while in Czech, the stressed syllable bears no prominence in any acoustic domain (Skarnitzl and Eriksson, 2017; Skarnitzl, 2018). This is not surprising given the fact that Czech is a language with stress fixed to the first syllable of a prosodic word and no contrastive function (*cf.* Cutler, 2005). Recent evidence also suggests that it may be suitable to talk about accentual groups rather than stress groups in Czech, with groups of words sometimes joined into one prosodic unit (Volín and Skarnitzl, 2020). In terms of rhythmic properties, English is a language characterized by large differences in syllable durations, with long stressed syllables and unstressed ones reduced in duration, as well as quality; traditionally English has been described as a stress-based language (Dauer, 1983). Czech has phonological vowel length, although phonologically long vowels are considerably less frequent in connected speech (see Volín, 2010, p. 45), and no systematic vowel reduction; traditionally Czech has been described as a syllable-based language, in spite of its syllabic complexity (Šturm and Lukeš, 2017; see also Dankovičová and Dellwo, 1999).

The monotonousness of Czech is evident not only in the weak to absent prominence contrasts, but particularly in the melodic domain. Compared to English, Czech intonation comes across as rather flat, and there seem to be two reasons for this. First, pitch range is significantly narrower in Czech: in a comparison of English and Czech professional newsreaders, Volín et al. (2015) report the 80-percentile range (i.e., the difference between the 90th and 10th percentile) to be 7.1 semitones (ST) for British females and 8.1 ST for British males, while the values in Czech were 5.2 ST and 6.1 ST, respectively. Melody seems to be even flatter in spontaneous speech: unpublished data show that in 56 out of 100 male speakers, 80-percentile range does not exceed 5 ST. The second reason for the monotonous impression is the fact that, apart from flatter melody, Czech is also characterized by longer prosodic phrases. In a comparison of British English and Czech radio newsreaders, Volín (2019) found that prosodic phrases in English were, on average, by nearly 40% longer in Czech than in English, with the mean lengths being 10.8 and 7.8 syllables, respectively; interestingly, there was no significant difference, however, when phrase length was expressed in words (4.6 in Czech and 4.5 in English). Volín also examined news reading by non-professional Czech speakers and reported even longer prosodic phrases, with the mean being 12.9 syllables and 5.4 words. It seems natural that segmenting utterances into a smaller number of longer prosodic phrases will further contribute to the perceived monotonousness of Czech intonation, as there are longer stretches of speech without salient melodic movements.

Research Questions

It was mentioned in "Charismatic Speech and Prosody" that the association of good speaker qualities with short prosodic phrases is, to the best of our knowledge (and also as confirmed by Niebuhr, November 2021, personal communication), based on rather anecdotal evidence. At the same time, observations of everyday Czech speech indicate that prosodic phrases are rather long in Czech, considerably longer than in English. Combining these two gaps in our knowledge, the general objective of the present study is to compare prosodic phrasing in good speakers of English and Czech, in the genre of public speaking.

Specifically, then, we are asking whether more monotonous prosody is also observed in good speakers of Czech, as compared with speakers of English, or whether the communicative demands of public speaking lead to a different prosodic behavior than observed in ordinary Czech speech. In other words, our goal is to compare genre- and language-specific tendencies in skilled speakers of Czech and English. Genre-based differences would be in agreement with the results of De Pijper and Sanderman (1994), who showed that professional speakers produced more prosodic cues and more salient ones than non-professional speakers. On the other hand, language-based differences may persist, with Czech and English speakers in the same genre still manifesting different prosodic tendencies. We will use measures of prosodic phrase length, articulation rate, pitch range, and melodic variability in the domain of the prosodic

phrase to describe good public speakers' prosodic phrasing in the two languages.

One of the contributions of this study is extending the scope of melodic measures. Apart from applying a still relatively new measure of melodic variability (see "Analyses" below), we focus not only on entire prosodic phrases, but also separately on their nuclear part (i.e., in the syllables which carry the nuclear tone of each phrase) and their pre-nuclear field (i.e., in all the syllables preceding the nuclear syllable). It is especially the analysis in the pre-nuclear field that makes this study different from the results reported by Hledíková (2019), on which this study is based. The motivation for including the pre-nuclear field comes from the informal observation that the flatness of Czech intonation may not be captured by traditional indicators of pitch range like standard deviation: the flatness may be limited only to the pre-nuclear field, with little melodic movement apart from the nuclear tone itself.

MATERIALS AND METHODS

Material

This study is based on TED Talks delivered in American English and Czech. In TED Talk events, speakers present a topic in an attractive and entertaining way to a general audience, in a relatively limited time span (typically between 15 and 20 min). In the first stage, we selected 15 speakers of each language; this selection was based on subjectively perceived speaker quality. We assumed that already the fact that a speaker was invited to a TED Talk event provides a certain guarantee of their speaking competence: in fact, TED Talks have been described in a recent study as "the pinnacle of public speaking" (Tsai, 2015).[1] However, to further ensure the high quality of the speakers' performance, we conducted an informal listening test with eight listeners per language. The respondents were played 30-s segments and asked to express their willingness to employ the speaker as their spokesperson on a 7-point scale. Ten TED Talk speakers of each language who received the highest mean score were chosen for subsequent analysis; this corresponded to six males and four females speakers in English, and nine males and one female speakers in Czech. Due to the gender imbalance, possible differences between female and male speakers will not be examined.

The recordings were divided into shorter segments of approximately 1 min and forced-aligned using P2FA for English (Yuan and Liberman, 2008) and Prague Labeller for Czech (Pollák et al., 2007). Five minutes of speech per speaker were selected; the first 2 min were not analyzed because the speaker may need time to "get started" and find his speaking style.

Analyses

Prosodic boundaries were labeled manually in Praat (Boersma and Weenink, 2019), using break indices 3 and 4 in line with ToBI conventions for minor and major prosodic breaks, respectively; any prosodic discontinuities were marked with a "p" (Beckman and Elam, 1997, p. 32). The labeling was carried out by the second author, and any uncertainties were settled in a joint analysis of the two authors. We also marked the syllable carrying nuclear stress in each phrase. A Praat script was then used to extract the following measures from the annotated data:

- number of syllables per prosodic phrase;
- number of words per prosodic phrase;
- articulation rate in syllables/s;
- f_0 standard deviation (SD) in each prosodic phrase in ST;
- Cumulative Slope Index (CSI) in each prosodic phrase in ST/syllable; CSI corresponds to the sum of absolute frequency differences between subsequent f_0 points divided by the number of syllables and thus captures melodic variability better than the more traditional measures, as it takes into account multiple melodic movements in a phrase (Hruška and Bořil, 2017);
- f_0 SD in the nuclear part of the phrase in ST; and
- f_0 SD in the pre-nuclear part of the phrase in ST.

Fundamental frequency was extracted using default settings for autocorrelation in Praat, only with the ceiling lowered (320 Hz for male and 450 Hz for female speakers). The extracted Pitch objects were smoothed using a 10-Hz filter to eliminate microprosodic fluctuations in f_0, interpolated, and converted into PitchTier objects which were used to measure the SD of f_0 in ST.

Linear mixed-effects (LME) models were constructed for the statistical analyses, using (R Core Team, 2017) and the *lme4* package (Bates et al., 2015). The prosodic measures listed above served as dependent variables. There were two fixed effects, LANGUAGE (English, Czech) and PROSODIC BREAK TYPE (BI4, BI3). Random effects included SPEAKER intercept (since speakers may differ in their prosodic behavior) and by-SPEAKER slope for the effect of PROSODIC BREAK TYPE (since speakers may differ in their realization of each type of prosodic break). Residual plots were visually inspected for deviations from normality and homoscedasticity. The significance of individual effects or interactions was tested by comparing the full model to a reduced model with the given factor or interaction excluded. We conducted Tukey post-hoc tests using the *multcomp* package (Hothorn et al., 2008) to test specific pairwise comparisons. Plots showing mean values of the measured variables and their confidence intervals were created using the *effects* package (Fox, 2003) and visualized with *ggplot2* (Wickham, 2009).

RESULTS

Results will be presented in two main sections focusing gradually on temporal measures (length of prosodic phrases and articulation rate) and on melodic measures. They will always be displayed separately for major (BI4) and minor (BI3) prosodic breaks. In addition, a separate analysis will be presented for phrases which feature no disfluencies (in other words, we are also

[1]This is not to suggest, however, that speakers on the TED platform do not differ in their speaking quality and charismatic impact, as confirmed, among others, by MacKrill et al. (2021).

interested in the prosodic behavior of fully developed and realized phrases); since disfluencies occurred prevalently in minor phrases (i.e., those ending with a minor prosodic break), only phrases ending in BI4 without disfluencies (marked BI4-d below) will be considered in these partial analyses. "Individual Prosodic Profiles" will be dedicated to assessing prosodic phrase length and melodic range in a single comparison, so as to focus on possible individual differences between our TED Talk speakers.

Temporal Aspects of Phrasing

Mean length of prosodic phrases in both languages, expressed in syllables and words, is shown in **Figure 1**. It is obvious that the syllable and word levels provide different results. As for syllables (shown on the left of the figure), both LANGUAGE [$\chi^2(1) = 7.80$, $p < 0.01$] and PROSODIC BREAK TYPE [$\chi^2(1) = 25.18$, $p < 0.0001$] turned out to be significant predictors of phrase length. Phrases delivered by our English speakers are shorter, on average, by 0.89 (± 0.25 standard errors) syllables than those uttered by the Czech speakers, and phrases ending in a stronger prosodic break are longer by 0.5–3.7 syllables (with differences between individual speakers). The figure also suggests an interaction between LANGUAGE and PROSODIC BREAK TYPE, which is indeed significant: $\chi^2(1) = 6.98$, $p < 0.01$. Tukey post-hoc tests reveal that it is the BI4 context (i.e., major prosodic breaks) where Czech and English speakers' phrases differ in length ($p < 0.001$).

When expressed in the number of words, the effect of LANGUAGE on prosodic phrase length is marginally significant [$\chi^2(1) = 2.88$, $p < 0.1$], with phrases in English longer by 0.43 word on average (± 0.13 standard errors). This is not surprising, because although Czech phrases are longer when expressed in syllables, the analytical English uses many short words with a grammatical function, such as articles or prepositions, as opposed to the synthetic Czech which uses inflections to express grammatical relationships. The results are similar for the effect of PROSODIC BREAK TYPE [$\chi^2(1) = 23.47$, $p < 0.0001$], with phrases ending in BI4 being longer by 0.5–1.6 words than those ending in BI3. Tukey post-hoc tests reveal no significant effects of LANGUAGE ($p > 0.1$).

When only phrases ending in the BI4-type break and containing no disfluencies (BI4-d; see above) are considered, the difference between Czech and English phrases becomes even more pronounced [$\chi^2(1) = 15.60$, $p < 0.0001$]. As shown in the left part of **Figure 2**, prosodic phrases in Czech are, on average, by 1.72 (± 0.37 standard errors) syllables longer than those in English. Since the difference (as compared with the red confidence intervals in the left part of **Figure 1**) is attributable mostly to Czech, it seems that many of the shorter phrases ending in BI4 in Czech can be accounted for by some kind of disfluency (hesitation, prolongation, etc.). When expressed in words (in the right part of **Figure 2**), the tendency for longer phrases in English than in Czech is not significant [$\chi^2(1) = 1.66$, $p > 0.1$].

Another perspective on the length of prosodic phrases in Czech and English is provided by the histogram in **Figure 3**; again, only BI4-d phrases are shown. We can see that phrases in both languages are most typically quite short, with the most frequent phrase length around four to five syllables. The number of phrases which extend beyond eight syllables declines steadily in both languages. However, beyond the ten-syllable mark, the counts are consistently higher for Czech than for English; in other words, longer phrases are much more likely to be found in Czech than English, and the difference between the two languages in the previous two figures is, at least partially, caused by the prevalence of such extra-long phrases in Czech.

Turning to articulation rate, we can see in **Figure 4** that the Czech speakers were slightly faster in their delivery than the English ones. When we consider all data (i.e., the relationships shown in red and blue), the LANGUAGE factor significantly affected

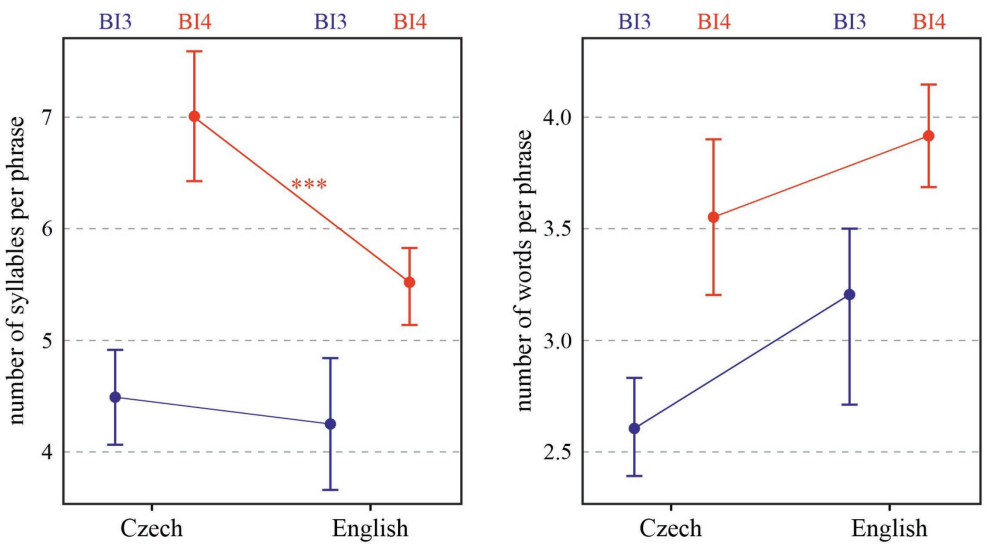


FIGURE 1 | Number of syllables (left) and words (right) per prosodic phrase depending on LANGUAGE and PROSODIC BREAK TYPE (BI3 in blue; BI4 in red). ***$p < 0.001$.

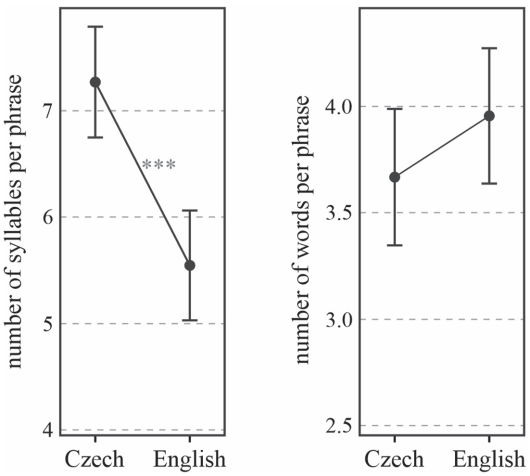


FIGURE 2 | Number of syllables (left) and words (right) per prosodic phrase depending on LANGUAGE, only in BI4 phrases and with no disfluencies (BI4-d). ***$p < 0.001$.

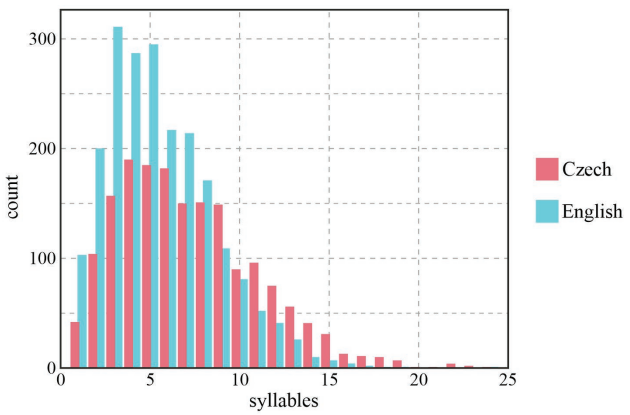


FIGURE 3 | Distribution of number of syllables per prosodic phrase depending on LANGUAGE, only in BI4 phrases and with no disfluencies (BI4-d).

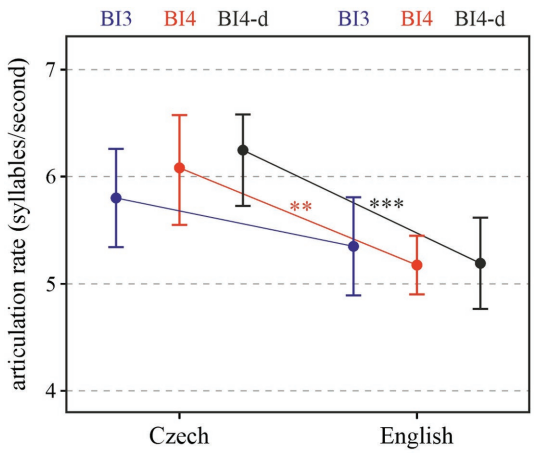


FIGURE 4 | Articulation rate depending on LANGUAGE and PROSODIC BREAK TYPE (BI3 in blue, BI4 in red, and BI4-d in black). **$p < 0.05$; ***$p < 0.001$.

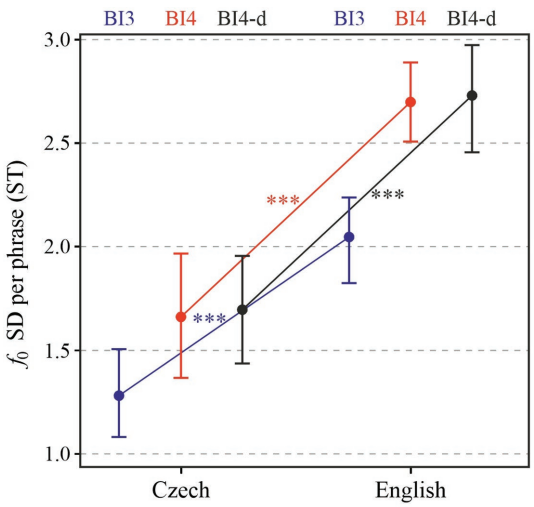


FIGURE 5 | Standard deviation of f_0 in each prosodic phrase depending on LANGUAGE and PROSODIC BREAK TYPE (BI3 in blue, BI4 in red, and BI4-d in black). ***$p < 0.001$.

articulation rate [$\chi^2(1) = 5.38$, $p < 0.05$], with Czech being faster by 0.76 (± 0.29 standard errors) syllables per second on average.

The effect of PROSODIC BREAK TYPE is not significant [$\chi^2(1) = 0.26$, $p > 0.6$]. When only BI4-d phrases (i.e., those featuring no disfluencies) are taken into account, the difference is significant [$\chi^2(1) = 8.69$, $p < 0.001$], with the difference between the two languages approaching one syllable per second: on average 0.96 (± 0.31 standard errors) syllables/s.

Melodic Patterning

First of all, we will examine melodic range in entire prosodic phrases. When expressed as the standard deviation of f_0 in each prosodic phrase, **Figure 5** shows that the speakers' LANGUAGE affected pitch range significantly [$\chi^2(1) = 15.98$, $p < 0.0001$]. On average, f_0 standard deviation in English was by 0.78 (± 0.15 standard errors) ST larger than in Czech. PROSODIC BREAK TYPE (BI4 or BI3) also significantly affected the standard deviation of f_0 in prosodic phrases [$\chi^2(1) = 26.74$, $p < 0.0001$], with differences between speakers ranging between 0.08 and 0.94 ST. Post-hoc tests, whose results are indicated using asterisks in the figure, indicate significant differences between f_0 standard deviation across languages: melodic range is higher in English than in Czech phrases ending in BI4 ($p < 0.0001$), as well as BI3 ($p < 0.001$). In addition, f_0 standard deviation in minor (BI3) phrases is, in each language, significantly lower than in major (BI4) phrases ($p < 0.0001$).

The results differ only very little when BI4-d phrases are considered (i.e., those BI4-type phrases which feature no disfluencies, shown in black in **Figure 5**): LANGUAGE again turned out to be a significant predictor of f_0 standard deviation [$\chi^2(1) = 19.48$, $p < 0.0001$], with mean values higher in English by 1.02 (± 0.19 standard errors) ST.

In **Figure 6**, we provide a slightly different perspective on the melodic behavior of our TED speakers' recordings. As explained in "Analyses", the Cumulative Slope Index (CSI) does not reflect only the range of f_0 values in a prosodic phrase, but rather captures potential multiple melodic movements within phrases, and may thus be a superior indicator of melodic variability. Overall, the effect of LANGUAGE on CSI is significant and similar in its magnitude to the standard deviation of f_0 reported above [$\chi^2(1) = 15.73$, $p < 0.0001$]. On average, melodic variability as expressed by CSI is by 1.62 ST/syllable (± 0.27 standard errors) higher in English than in Czech. The analysis of the effect of PROSODIC BREAK TYPE yielded a singular fit and could not be carried out. The interaction between LANGUAGE and PROSODIC BREAK TYPE (BI4 and BI3) turned out to be significant [$\chi^2(1) = 9.35$, $p < 0.01$]. As for phrases ending in a BI4 break and containing no disfluencies, LANGUAGE significantly affected CSI values [$\chi^2(1) = 24.29$, $p < 0.0001$], with mean CSI higher by 1.99 (± 0.30 standard errors) ST/syllable in English than in Czech. Tukey post-hoc test results reveal that CSI values do not significantly differ between major (BI4) and minor (BI3) phrases in Czech, while in English CSI is significantly lower in minor than in major phrases ($p < 0.0001$). The main difference in melodic variability between the two languages therefore consists in the major prosodic phrases, regardless of the presence or absence of disfluencies ($p < 0.0001$).

Finally, as we mentioned at the end of "Research Questions", it may be illustrative to consider melodic variability separately in the nuclear and pre-nuclear field, because the divergent impressions of English and Czech melodic patterning may be more strongly related to one or the other, as compared with the entire prosodic phrase. The results of standard deviation of f_0 in the nuclear portion of prosodic phrases are shown in **Figure 7**. The effect of LANGUAGE is significant [$\chi^2(1) = 15.97$, $p < 0.0001$], with f_0 standard deviation higher by 0.67 (± 0.13 standard errors) ST on average. The effect of PROSODIC BREAK TYPE also turned out to be significant [$\chi^2(1) = 27.99$, $p < 0.0001$]. Tukey post-hoc tests show that the difference between BI3 and BI4 is significant both in Czech and in English ($p < 0.0001$) and that cross-language comparisons by prosodic break type are also significant ($p < 0.0001$ for BI4 and $p < 0.01$ for BI3). When only BI4-d phrases are considered, the results for f_0 standard deviation in the nuclear portion are similar [$\chi^2(1) = 17.83$, $p < 0.0001$], with mean values being higher by 1 (± 0.2 standard errors) ST.

Results for standard deviation of f_0 in the pre-nuclear field of prosodic phrases are shown in **Figure 8**. The effect of LANGUAGE is significant [$\chi^2(1) = 18.03$, $p < 0.0001$], with f_0 standard deviation higher by 0.7 (± 0.12 standard errors) ST

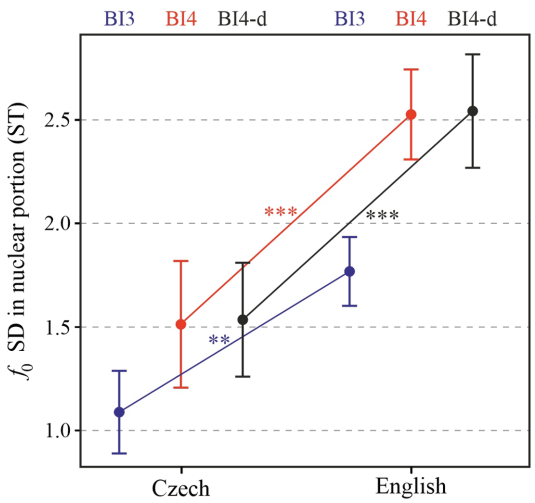


FIGURE 7 | Standard deviation of f_0 in the nuclear portion of each prosodic phrase, depending on LANGUAGE and PROSODIC BREAK TYPE (BI3 in blue, BI4 in red, and BI4-d in black). $^{**}p < 0.05$; $^{***}p < 0.001$.

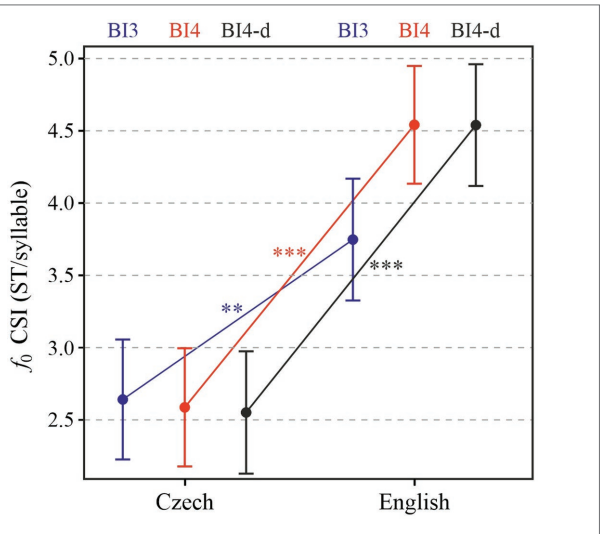


FIGURE 6 | Cumulative slope index (CSI) of f_0 in each prosodic phrase depending on LANGUAGE and PROSODIC BREAK TYPE (BI3 in blue, BI4 in red, and BI4-d in black). $^{**}p < 0.05$; $^{***}p < 0.001$.

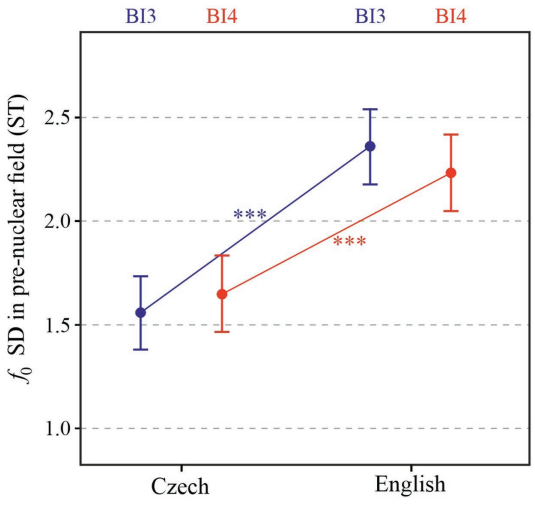


FIGURE 8 | Standard deviation of f_0 in the pre-nuclear field of each prosodic phrase depending on LANGUAGE and PROSODIC BREAK TYPE (BI3 in blue, BI4 in red, and BI4-d in black). $^{***}p < 0.001$.

on average. Unlike in previous analyses, the effect of PROSODIC BREAK TYPE is not significant [$\chi^2(1) = 0.08$, $p > 0.5$]; however, the interaction between the two factors does reach significance [$\chi^2(1) = 4.78$, $p < 0.05$]. As indicated using the asterisks in the figure, post-hoc comparisons confirm significant differences between f_0 standard deviation in Czech and English for both BI3-type and BI4-type phrases ($p < 0.0001$).

Individual Prosodic Profiles

The aim of this section is to provide a glimpse at individual variability between our speakers by placing length of prosodic phrases and melodic variability (CSI) next to each other. The comparison is shown using boxplots in **Figure 9** (Note, first, that only BI4-d phrases are shown and second, that the scales of the two variables are identical when outliers are included).

The average tendencies reported in the previous sections—longer phrase durations (always on the left for each speaker) and lower melodic variability (on the right) in Czech than in English—are clearly visible from the plot. It is not surprising that the mean values conceal considerable individual variability. In terms of phrase length, speaker CZ15's distribution is quite similar to that of most English speakers. On the other hand, phrases of speakers like CZ02 or CZ12 are on average twice as long as those of speaker CZ15. Turning to melodic variability, we can see that speakers EN07 and EN15 differ very little from some of the Czech speakers. In Czech, speaker CZ10's distribution is extremely narrow, with all prosodic phrases manifesting a CSI of f_0 lower than 5 ST/syllable.

It is also interesting to consider the highest values of both variables. For phrase length, while all phrases longer than 15 syllables correspond to outliers in English (represented by individual dots in **Figure 9**), and there are only nine of such phrases, one half of the Czech speakers' "normal" values (i.e., those within the boxplot whiskers) exceed 15 syllables, and there are 50 phrases in the Czech material exceeding 15 syllables in length. For melodic variability, it is obvious at first sight that the English material features more outliers, and their values reach higher values.

DISCUSSION

This study has focused on prosodic patterning in good public speakers active on the TED Talk platform, with the prosodic phrase as the domain of analysis. We analyzed ten speakers of Czech and ten of American English, two languages which differ considerably in their prosodic patterning. Our objective was to find whether the language-specific differences persist also in this genre—public speaking in an entertaining manner in front of an audience—or whether genre-specific requirements will bring speakers of the languages closer to each other in terms of their temporal and melodic behavior.

Our results indicate differences between Czech and English public speakers in both the temporal and melodic domain. As for the former, the length of prosodic phrases, when expressed in the number of syllables, differs especially in major prosodic phrases (those followed by a BI4 break). However, the phrases were considerably shorter in both languages than those reported by Volín (2019) for news reading on the radio, as shown by the comparison in **Table 1**. The prevalence of rather short phrases was clear in the histogram in **Figure 3**, and it is clear that the genre exerts a strong effect on the prosodic segmentation of speech: good speakers in both Czech and English use shorter prosodic phrases, they divide their flow of speech more often by prosodic breaks to achieve a better effect on the audience. Some studies also indicate that shorter prosodic phrases may be characteristic of skilled public speakers (Strangert, 2005). This is not surprising, as a more structured speech makes the process of understanding the message easier, and shorter segments are easier to process. Short phrases often appeared in contexts when our public speakers used emphasis; they frequently realized such words in separate prosodic phrases, as shown by the following examples from our data:

1. It turned out to be | shame.
2. A říká se tomu | exotika (*and it's called* | *exotica*).

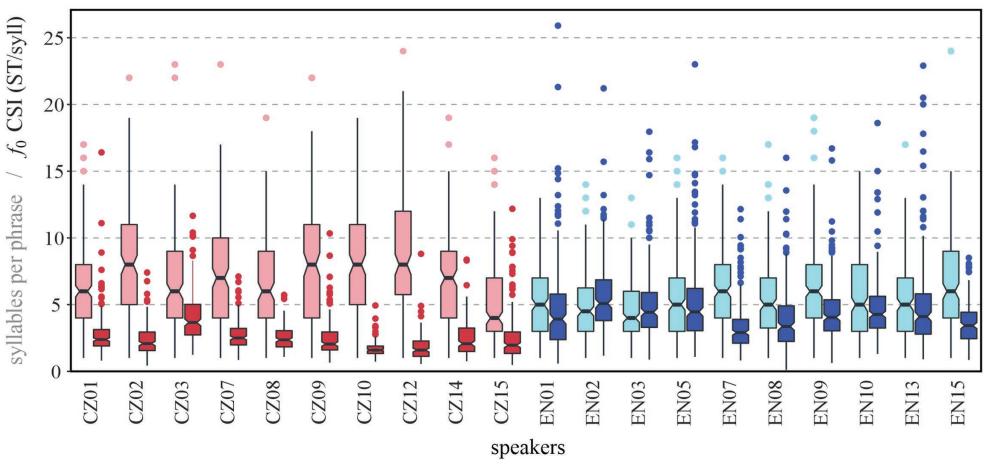


FIGURE 9 | Prosodic phrase length in syllables (boxplots on the left in lighter colors) and cumulative slope index (CSI) of f_0 (boxplots on the right in darker colors) for individual speakers in BI4-d phrases.

TABLE 1 | Mean length of prosodic phrases (in syllables) in English and Czech in news reading (Volín, 2019) and in TED speakers (this study; only major phrases with no disfluencies are included).

	English	Czech
Radio newsreaders	7.8	10.8
TED speakers	5.5	7.2

TABLE 2 | 80-percentile range of f_0 (in semitones) in English and Czech in news reading (Volín et al., 2015; male and female values are averaged here) and in TED speakers (this study; only major phrases with no disfluencies are included).

	English	Czech
Radio newsreaders	7.60	5.65
TED speakers	7.05	4.48

Let us consider the question of language- vs. genre-specific tendencies mentioned above. As shown in **Table 1**, phrases are longer in Czech than in English by 31% in our material, while in Volín's data, this difference amounts to 38%. It appears, therefore, that phrasing is influenced by both genre (resulting in shorter phrases in public speaking than in news reading) and language (with phrases longer in Czech than in English).

As for melodic patterning, the results of this study point to lower values of pitch range (standard deviation of f_0) and overall melodic variability (cumulative slope index, CSI) in Czech than in English, thus lending support to previous studies (especially Volín et al., 2015), as well as informal observations. To make our results comparable with those of Volín et al. (2015), we also calculated the 80-percentile ranges of the f_0 data; the comparison is shown in **Table 2**. It can be seen that pitch range is narrower in our data when compared with news reading, although the difference is quite small: it amounts to a little over one quartertone in English and slightly exceeds one semitone in Czech. In melodic patterning, our data therefore suggest that language-specific tendencies exert a greater influence on the speakers' pitch range than the genre.

Unfortunately, it is not possible to compare our results with studies reported in "Charismatic Speech and Prosody" which targeted pitch range: none of these report absolute values, only correlation coefficients relating pitch range to perceived charisma (or another measure like the number of subscriptions).

In this study, we adopted a closer look at melodic variation within a prosodic phrase by considering not only the entire phrase as a unit, but also dividing it into the nuclear part (carrying the nuclear tone) and the pre-nuclear field. Specifically, we hypothesized that the impression of flat melody in Czech may be due to little variation in the pre-nuclear field, especially when combined with the tendency to produce longer prosodic phrases in Czech. However, our results do not suggest a difference in the standard deviation of f_0 in the pre-nuclear or nuclear portion of prosodic phrases (see **Figures 7, 8**): in both Czech and English, f_0 standard deviation appears to be quite similar in both portions of prosodic phrases. In fact, it seems to be English where melodic range is smaller in the pre-nuclear field than in the nuclear part. It would be interesting, however, to conduct comparable analyses on spontaneous dialogs or other speaking styles in Czech and English: the melodic difference between the pre-nuclear and nuclear field may become more pronounced there.

There are some minor findings which are worth commenting on. First of all, it should be noted that the length of prosodic phrases does not significantly differ between the two examined languages when it is expressed in words. The same result was reported by Volín (2019), as already mentioned in "Prosodic Patterns in English and Czech". This seems to be due to the different morphological type of Czech and English, with Czech using inflections where English uses individual words.

In the current study, we analyzed separately phrases ending in a weaker, BI3-type of break (also referred to as minor phrases above) and those ending in a stronger, BI4-type of break (major phrases). Our results showed a significant effect of PROSODIC BREAK TYPE, with minor phrases being shorter and having a narrower pitch range than major phrases in both languages. In addition, we identified those phrases which included a disfluency (most frequently these were minor phrases); subsequently, we presented results for phrases without any disfluencies (marked BI4-d). One of the reasons for this was our expectation that these "full-fledged" phrases are what is typically analyzed in most other studies; that is, why the above comparisons with studies by Volín and colleagues featured only these BI4-d phrases.

Finally, articulation rate turned out to differ between our Czech and American English speakers; interestingly, the mean values correspond quite closely to those reported in literature, even for other speaking styles or genres. The mean articulation rate (in BI4-d phrases only) in our Czech data is 6.2 syllables per second, similar to the 6.1 syll/s reported by Veroňková and Poukarová (2017) for Czech newsreaders. Our American English speakers' mean articulation rate was 5.2 syll/s, while that in the read speech examined by Baese-Berk and Morrill (2015) was 4.9 syll/s. Note that while these are relatively fast speech rates, good quality speaking has been associated with faster speaking (see "Charismatic Speech and Prosody").

The comparison of individual tendencies in "Individual Prosodic Profiles" provided a useful perspective on our data, and it is crucial for the concept of a good or charismatic speaker. All our speakers were chosen by listeners from a larger dataset as high-quality speakers (*cf.* "Material" for speaker selection). To take but one example, given that speaker CZ10's melody is flattest and his phrases belong among the longest, it seems obvious that being a "good speaker" involves a constellation of a number of factors which may partially compensate for each other. Based on our current knowledge (see "Charismatic Speech and Prosody"), temporal and melodic variability appear to be crucial but not the only components of such a constellation, and speaker charisma should be treated as a multidimensional phenomenon.

To conclude, the current study has confirmed previous studies and informal observations that the Czech language is prosodically more monotonous than English but extended them by analyzing skilled speakers who have been selected for their high-quality delivery. In our TED speakers, Czech prosodic phrases are approximately 30% longer than English ones and, at the same time, variation in the melodic domain is much lower, with f_0 standard deviation smaller by one semitone and melodic variability (expressed by CSI) by nearly two semitones smaller in Czech than English. It is not surprising, then, that such a combination results in the perception of monotonousness or flatness; naturally, this will be particularly salient to listeners whose native language manifests greater melodic variability. There seems to be a smaller difference between the two languages in phrase length than in other genres, showing a combined effect of language and genre, but pitch range remained mostly language-specific. The natural follow-up to this study would be moving from skilled, charismatic speakers to ordinary ones and their prosodic behavior in everyday conversations.

The implications of our study extend to speakers who strive for high-quality performance. It is clear that prosodic rendition of one's speech—especially the length of prosodic phrases, melodic variability, and speech rate—is a vital aspect of speech performance. Given the fact that these prosodic aspects of speech are, to a considerable extent, language-specific, speakers should take special care when they deliver a speech or give a presentation in a foreign language: what may sound like charismatic speaking in one language may sound flat and disinterested or, on the other hand, affected and insincere in another.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

FUNDING

This study was supported by the Czech Science Foundation project "Prosodic Phrase in Current Spoken Czech: Meaning, Balance, Stochastic Patterns" (no. GA 21-14758S) and from the European Regional Development Fund-Project "Creativity and Adaptability as Conditions of the Success of Europe in an Interrelated World" (no. CZ.02.1.01/0.0/0.0/16_019/0000734).

REFERENCES

Amon, I. (2016). *Die Macht der Stimme*. Munich: Redline.

Andreeva, B., Demenko, G., Wolska, M., Möbius, B., Zimmerer, F., Jügler, J., et al. (2014). "Comparison of pitch range and pitch variation in Slavic and Germanic languages", in *Proceedings of Speech Prosody*; May 20–23, 2014, Dublin, Ireland, 776–780.

Antonakis, J., Fenley, M., and Liechti, S. (2011). Can charisma be taught? Tests of two interventions. *Acad. Manag. Learn. Educ.* 10, 374–396. doi: 10.5465/amle.2010.0012

Awamleh, R., and Gardner, W. L. (1999). Perceptions of leader charisma and effectiveness: The effects of vision content, delivery, and organizational performance. *Leadership Quart.* 10, 345–373. doi: 10.1016/S1048-9843(99)00022-3

Baese-Berk, M. M., and Morrill, T. H. (2015). Speaking rate consistency in native and non-native speakers of English. *J. Acoust. Soc. Am.* 138, EL223–EL228. doi: 10.1121/1.4929622

Barker, A. (2011). *Improve Your Communication Skills. 2nd Edn*. London: Kogan Page.

Barker, A. (2019). *Improve Your Communication Skills: How to Build Trust, Be Heard and Communicate With Confidence. 5th Edn*. London: Kogan Page.

Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* 67, 1–48. doi: 10.18637/jss.v067.i01

Beckman, M. E., and Elam, G. A. (1997). Guidelines for ToBI labelling (version 3.0). Ohio State University: The Ohio State University Research Foundation. Available at: https://docplayer.net/8594926-Guidelines-for-tobi-labelling-version-3-0-march-1997-by-mary-e-beckman-and-gayle-ayers-elam.html (Accessed August 2008).

Berger, S., Niebuhr, O., and Zellers, M. (2019). "A preliminary study of charismatic speech on YouTube: correlating prosodic variation with counts of subscribers, views and likes," in *Proceedings of Inter Speech 2019*; September 15–19, 2019, Graz, Austria, 1761–1765.

Biadsy, F., Rosenberg, A., Carlson, R., Hirschberg, J., and Strangert, E. (2008). "A cross-cultural comparison of American, Palestinian, and Swedish perception of charismatic speech," in *Proceedings of Speech Prosody*; May 6–9, 2008, Campinas, Brazil, 579–582.

Boersma, P., and Weenink, D. (2019). Praat: doing phonetics by computer. Version 6.1. Available at: http://www.praat.org (Accessed July 14, 2019).

Bolkan, S., and Goodboy, A. K. (2014). Communicating charisma in instructional settings: indicators and effects of charismatic teaching. *Coll. Teach.* 62, 136–142. doi: 10.1080/87567555.2014.956039

Caspi, A., Bogler, R., and Tzuman, O. (2019). "Judging a book by its cover": The dominance of delivery over content when perceiving charisma. *Group Organ. Manage.* 44, 1067–1098. doi: 10.1177/1059601119835982

Chen, A., Gussenhoven, C., and Rietveld, T. (2004). Language-specificity in the perception of paralinguistic intonational meaning. *Lang. Speech* 47, 311–349. doi: 10.1177/00238309040470040101

Cresti, E. (2018). The illocution-prosody relationship and the information pattern in spontaneous speech according to the language into act theory (L-AcT). *Linguistik Online* 88, 33–62. doi: 10.13092/lo.88.4189

Cutler, A. (2005). "Lexical stress," in *The Handbook of Speech Perception*. eds. D. B. Pisoni and R. E. Remez (United Kingdom: Blackwell Publishing), 264–289.

Dankovičová, J., and Dellwo, V. (1999). "Czech speech rhythm and the rhythm class hypothesis," in *Proceedings of the 16th International Congress of Phonetic Sciences*; August 1–7, 1999, San Francisco, United States, 1241–1244.

Dauer, R. M. (1983). Stress-timing and syllable-timing reanalyzed. *J. Phonetics* 11, 51–62. doi: 10.1016/S0095-4470(19)30776-4

De Jong, N., and Mora, J. (2019). Does having good articulatory skills lead to more fluent speech in first and second languages? *Stud. Sec. Lang. Acq.* 41, 227–239. doi: 10.1017/S0272263117000389

De Pijper, J. R., and Sanderman, A. A. (1994). On the perceptual strength of prosodic boundaries and its relation to suprasegmental cues. *J. Acoust. Soc. Am.* 96, 2037–2047. doi: 10.1121/1.410145

D'Errico, F., Signorello, R., Demolin, D., and Poggi, I. (2013). "The perception of charisma from voice. A crosscultural study," in *Proceedings of Humaine Association Conference on Affective Computing and Intelligent Interaction*; September 2–5, 2013, Geneva, Switzerland, 552–557.

Derwing, T. M., and Rossiter, M. J. (2003). The effects of pronunciation instruction on the accuracy, fluency, and complexity of L2 accented speech. *Appl. Lang. Learn.* 13, 1–17.

Eriksson, A., and Heldner, M. (2015). "The acoustics of word stress in English as a function of stress level and speaking style," in *Proceedings of Inter Speech 2015*; September 6–10, 2015, Dresden, Germany, 41–45.

Fanelli, A., and Misangyi, V. F. (2006). Bringing out charisma: CEO charisma and external stakeholders. *Acad. Manag. Rev.* 31, 1049–1061. doi: 10.2307/20159265

Féry, C. (2017). *Intonation and Prosodic Structure*. Cambridge: Cambridge University Press.

Fischer, K., Niebuhr, O., Jensen, L. C., and Bodenhagen, L. (2019). Speech melody matters – how robots profit from using charismatic speech. *ACM Trans. Hum. Robot Int.* 9, 1–21. doi: 10.1145/3344274

Fox, J. (2003). Effect displays in R for generalised linear models. *J. Stat. Softw.* 8, 1–27. doi: 10.18637/jss.v008.i15

Frazier, L., Carlson, K., and Clifton, C. Jr. (2006). Prosodic phrasing is central to language comprehension. *Trends Cogn. Sci.* 10, 244–249. doi: 10.1016/j.tics.2006.04.002

Gélinas-Chebat, C., Chebat, J. C., and Vaninsky, A. (1996). Voice and advertising: effects of intonation and intensity of voice on source credibility, attitudes and the intent to buy. *Percept. Motor Skills* 83, 243–262. doi: 10.2466/pms.1996.83.1.243

Gendron, A. (2017). The call to jihad: charismatic preachers and the internet. *Stud. Confl. Terror.* 40, 44–61. doi: 10.1080/1057610X.2016.1157406

Gilner, L. (2014). A deconstruction of public speaking aimed at the development of presentation skills. *IATEFL Pro. Special Int. Group News.* 51, 24–35.

Grabe, E., Rosner, B. S., García-Albea, J. E., and Zhou, X. (2003). Perception of English intonation by English, Spanish, and Chinese listeners. *Lang. Speech* 46, 375–401. doi: 10.1177/00238309030460040201

Hledíková, H. (2019). Prosodic Phrasing in Good Speakers in English and Czech. Bachelor's thesis. Prague: Charles University.

Holladay, S. J., and Coombs, W. T. (1994). Speaking of visions and visions being spoken an exploration of the effects of content and delivery on perceptions of leader charisma. *Manage. Commun. Q.* 8, 165–189. doi: 10.1177/0893318994008002002

Hothorn, T., Bretz, F., and Westfall, P. (2008). Simultaneous inference in general parametric models. *Biom. J.* 50, 346–363. doi: 10.1002/bimj.200810425

Hruška, R., and Bořil, T. (2017). Temporal variability of fundamental frequency contours. *Acta Univ. Carolinae – Philo.* 2017, 35–44. doi: 10.14712/24646830.2017.31

Jun, S.-A. (2003). Prosodic phrasing and attachment preferences. *J. Psycholinguist. Res.* 32, 219–249. doi: 10.1023/a:1022452408944

Krivokapić, J. (2007). Prosodic planning: effects of phrasal length and complexity on pause duration. *J. Phon.* 35, 162–179. doi: 10.1016/j.wocn.2006.04.001

Leonard, L. B. (1974). The role of intonation in the recall of various linguistic stimuli. *Lang. Speech* 16, 327–335. doi: 10.1177/002383097301600403

Lower Mekong Initiative (2017). *Professional Communication Skills for Leaders: Core Curriculum*. United States: United States Department of State.

Lucas, S. E. (2015). *The Art of Public Speaking. 12th Edn*. New York: McGraw-Hill Education.

MacKrill, K., Silvester, C., Pennebaker, J. W., and Petrie, K. J. (2021). What makes an idea worth spreading? Language markers of popularity in TED talks by academics and other speakers. *J. Assoc. Inf. Sci. Tech.* 72, 1028–1038. doi: 10.1002/asi.24471

McDougall, K., and Duckworth, M. (2017). Profiling fluency: An analysis of individual variation in disfluencies in adult males. *Speech Commun.* 95, 16–27. doi: 10.1016/j.specom.2017.10.001

Mennen, I., Schaeffler, F., and Docherty, G. (2012). Cross-language differences in fundamental frequency range: A comparison of English and German. *J. Acoust. Soc. Am.* 131, 2249–2260. doi: 10.1121/1.3681950

Michalsky, J., and Niebuhr, O. (2019). Myth busted? Challenging what we think we know about charismatic speech. *Acta Univ. Carolinae – Philo.* 2019, 27–56. doi: 10.14712/24646830.2019.17

Mixdorff, H., Niebuhr, O., and Hönemann, A. (2018). "Model-based prosodic analysis of charismatic speech," in *Proceedings of Speech Prosody 2018*; June 13–16, 2018, Poznań, Poland, 814–818.

Niebuhr, O. (2017). "Clear speech – mere speech? How segmental and prosodic speech reduction shape the impression that speakers create on listeners," in *Proceedings of Inter Speech 2017*; August 20–24, 2017, Stockholm, Sweden, 894–898.

Niebuhr, O., and Gonzalez, S. (2019). Do sound segments contribute to sounding charismatic? Evidence from acoustic vowel space analyses of Steve jobs and mark Zuckerberg. *Int. J. Acoust. Vib.* 24, 343–355.

Niebuhr, O., and Neitsch, J. (2020). "Digital rhetoric 2.0: How to train charismatic speaking with speech-melody visualization software," in *Speech and Computer. SPECOM 2020. Lecture Notes in Computer Science. Vol. 12335*. eds. A. Karpov and R. Potapova (Cham, Germany: Springer), 357–368.

Niebuhr, O., Skarnitzl, R., and Tylečková, L. (2018a). "The acoustic fingerprint of a charismatic voice – initial evidence from correlations between long-term spectral features and listener ratings," in *Proceedings of Speech Prosody 2018*; June 13–16, 2018, Poznań, Poland.

Niebuhr, O., Tegtmeier, S., and Brem, A. (2017). Advancing research and practice in entrepreneurship through speech analysis – from descriptive rhetorical terms to phonetically informed acoustic charisma metrics. *J. Speech Sci.* 6, 3–26. doi: 10.20396/joss.v6i1.14983

Niebuhr, O., Tegtmeier, S., and Schweisfurth, T. (2019). Female speakers benefit more than male speakers from prosodic charisma training – A before-after analysis of 12-weeks and 4-h courses. *Front. Commun.* 4:12. doi: 10.3389/fcomm.2019.00012

Niebuhr, O., Thumm, J., and Michalsky, J. (2018b). "Shapes and timing in charismatic speech – evidence from sounds and melodies," in *Proceedings of Speech Prosody 2018*; June 13–16, 2018, Poznań, Poland.

Niebuhr, O., Voße, J., and Brem, A. (2016). What makes a charismatic speaker? A computer-based acoustic prosodic analysis of Steve jobs tone of voice. *Comput. Hum. Behav.* 64, 366–382. doi: 10.1016/j.chb.2016.06.059

Nooteboom, S. G. (1995). "How far do we look ahead while speaking?" in *International Congress of Phonetic Sciences. Vol. 4*. eds. K. Ellenius and P. Branderud (Stockholm, Sweden: ISPhS), 578–581.

O'Connell, D. C., Turner, E. A., and Onuska, L. A. (1968). Intonation, grammatical structure, and contextual association in free recall. *J. Verb. Learn. Verb. Be.* 7, 110–116. doi: 10.1016/S0022-5371(68)80173-2

Pawley, A., and Syder, F. (2000). "The one clause at a time hypothesis," in *Perspectives on Fluency*. ed. H. Riggenbach (Ann Arbor, MI: University of Michigan Press), 163–191.

Petrone, C., Fuchs, S., and Krivokapić, J. (2011). "Consequences of working memory differences and phrasal length on pause duration and fundamental frequency," in *Proceedings of the 9th International Seminar on Speech Production (ISSP)*; June 20–23, 2011, Montréal, Canada.

Pollák, P., Volín, J., and Skarnitzl, R. (2007). "HMM-based phonetic segmentation in Praat environment," in *Proceedings of SPECOM 2007*; October 15–18, 2007, Moscow, Russia, 537–541.

R Core Team (2017). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

Reeves, C., Schmauder, A. R., and Morris, R. K. (2000). Stress grouping improves performance on an immediate serial list recall task. *J. Exp. Psychol. Learn. Mem. Cogn.* 26, 1638–1654. doi: 10.1037//0278-7393.26.6.1638

Rosenberg, A., and Hirschberg, J. (2009). Charisma perception from text and speech. *Speech Commun.* 51, 640–655. doi: 10.1016/j.specom.2008.11.001

Shattuck-Hufnagel, S., and Turk, A. E. (1996). A prosody tutorial for investigators of auditory sentence processing. *J. Psycholinguist. Res.* 25, 193–247. doi: 10.1007/BF01708572

Sheafer, T. (2008). Charismatic communication skill, media legitimacy, and electoral success. *J. Polit. Mark.* 7, 1–24. doi: 10.1080/15377850802063983

Skarnitzl, R. (2018). Fonetická realizace slovního přízvuku u delších slov v češtině [The phonetic realization of lexical stress in longer words in Czech]. *Slovo a slovesnost* 79, 199–216.

Skarnitzl, R., and Eriksson, A. (2017). "The acoustics of word stress in Czech as a function of speaking style," in *Proceedings of Inter Speech 2017*; August 20–24, 2017, Stockholm, Sweden, 3221–3225.

Strangert, E. (2005). "Prosody in public speech: analyses of a news announcement and a political interview," in *Proceedings of Inter Speech 2005*; September 4–8, 2005, Lisbon, Portugal, 3401–3404.

Strangert, E., and Gustafson, J. (2008). "What makes a good speaker? Subject ratings, acoustic measurements and perceptual evaluations," in *Proceedings of Inter Speech 2008*; September 22–26, 2008, Brisbane, Australia, 1688–1691.

Sturges, P. T., and Martin, J. G. (1974). Rhythmic structure in auditory pattern perception and immediate memory. *J. Exp. Psych.* 102, 377–383. doi: 10.1037/h0035866

Šturm, P., and Lukeš, D. (2017). Fonotaktická analýza obsahu slabik na okrajích českých slov v mluvené a psané řeči [A phonotactic analysis of the content of syllables on word boundaries in spoken and written Czech texts]. *Slovo a slovesnost* 78, 99–118.

Suzuki, S., Kormos, J., and Uchihara, T. (2021). The relationship between utterance and perceived fluency: A meta-analysis of correlational studies. *Mod. Lang. J.* 105, 435–463. doi: 10.1111/modl.12706

Swets, B., Desmet, T., Hambrick, D. Z., and Ferreira, F. (2007). The role of working memory in syntactic ambiguity resolution: A psychometric approach. *J. Exp. Ps.: General* 136, 64–81. doi: 10.1037/0096-3445.136.1.64

Touati, P. (1993). "Prosodic aspects of political rhetoric," in *Proceedings of the ESCA Workshop on Prosody*. September 27–29, 1993; Lund, Sweden; 168–171.

Towler, A. J. (2003). Effects of charismatic influence training on attitudes, behavior, and performance. *Pers. Psychol.* 56, 363–381. doi: 10.1111/j.1744-6570.2003.tb00154.x

Tsai, T. J. (2015). "Are you TED Talk material? Comparing prosody in professors and TED speakers," in in *Proceedings of Inter Speech 2015*; September 6–10, 2015, Dresden, Germany, 2534–2538.

Veroňková, J., and Poukarová, P. (2017). The relation between subjective and objective assessment of speaking rate in Czech radio newsreaders. *Acta Univ. Carolinae – Philo.* 2017, 95–107. doi: 10.14712/24646830.2017.36

Volín, J. (2010). "Fonetika a fonologie," in *Mluvnice Současné Češtiny [Grammar of contemporary Czech]*. ed. V. Cvrček (Praha: Karolinum), 35–64.

Volín, J. (2019). The size of prosodic phrases in native and foreign-accented read-out monologues. *Acta Univ. Carolinae – Philo.* 2019, 145–158. doi: 10.14712/24646830.2019.23

Volín, J., Poesová, K., and Weingartová, L. (2015). Speech melody properties in English, Czech and Czech English: reference and interference. *Res. Lang.* 13, 107–123. doi: 10.1515/rela-2015-0018

Volín, J., and Skarnitzl, R. (2020). "Accent-groups vs. stress-groups in Czech clear and conversational speech," in *Proceedings of Speech Prosody 2020*; May 25–28, 2020, Tokyo, Japan, 695–699.

Weber, M. (1947). *The Theory of Social and Economic Organization*. New York: The Free Press of Glencoe.

Wickham, H. (2009). *ggplot2: Elegant Graphics for Data Analysis*. New York: Springer.

Yuan, J., and Liberman, M. (2008). "Speaker identification on the SCOTUS corpus," in *Proceedings of Acoustics 2008*; June 29–July 4, 2008, Paris, France, 5687–5690

Zurif, E. B., and Mendelsohn, M. (1972). Hemispheric specialization for the perception of speech sounds: The influence of intonation and structure. *Percept. Psychophys.* 11, 329–332. doi: 10.3758/BF03206262

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The handling editor declared a past co-authorship with one of the authors RS.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Skarnitzl and Hledíková. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Psychology

ORIGINAL RESEARCH
published: 29 March 2022
doi: 10.3389/fpsyg.2022.855063

Research on Bond Participants' Emotion Reactions Toward the Internet News in China's Bond Market

Wei Zhang[1], *Jun Wang*[1]* *and Mu Tong*[2]*

[1] School of Economic Information Engineering, Southwestern University of Finance and Economics, Chengdu, China,
[2] Collaborative Innovation Center of Financial Security, Southwestern University of Finance and Economics, Chengdu, China

The literature has widely studied the market response to the financial news or events but mainly focused on the stock market. This article associates the concept of internet news with the bond market response and attempts to examine how credit rating agencies (CRAs) and bond investors, two important bond participants, react to financial news on the internet with a range of multiply regressions. Our empirical study leads to several findings. First, CRAs tend to ignore the warnings of financial news on the internet, whereas bond investors strongly react to such news. Second, there is an asymmetry in bond investors' reactions to good news compared to bad news, with investors being more sensitive to bad news. Third, there is heterogeneity in the psychological reaction where bond investors do not react to the news about central state-owned enterprises (SOEs) but to the news about other enterprises. Finally, there is an asymmetric response driven by news timeliness that bond investors are more sensitive to the latest news articles than old ones. Overall, our study confirms the existence of psychological reactions to the financial news on the internet in China's bond market, which has significance for keeping bond market participants from overreacting or underreacting to market news.

Keywords: emotion reaction, financial news, bond market, bond investor, credit rating agency

OPEN ACCESS

Edited by:
Anna Esposito,
University of Campania "Luigi Vanvitelli", Italy

Reviewed by:
Jakub Horák,
Institute of Technology and Business in České Budějovice, Czechia
Chenguel Bechir,
University of Kairouan, Tunisia

***Correspondence:**
Jun Wang
wangjun1987@swufe.edu.cn
Mu Tong
tongmu_swufe@outlook.com

Specialty section:
This article was submitted to Emotion Science, a section of the journal Frontiers in Psychology

Received: *14 January 2022*
Accepted: *22 February 2022*
Published: *29 March 2022*

Citation:
Zhang W, Wang J and Tong M (2022) Research on Bond Participants' Emotion Reactions Toward the Internet News in China's Bond Market. Front. Psychol. 13:855063. doi: 10.3389/fpsyg.2022.855063

INTRODUCTION

China has been the second-largest bond market globally, with an annual issuance of 57.3 trillion RMB in 2020, significantly benefiting the Chinese economy to achieve rapid growth. Despite the fast development, the information disclosure in China's bond market is still relatively immature. For example, managers tend to disclose some good information but delay the announcement about bad information (Tang et al., 2018). There is limited information available about the financial market, making the market more susceptible to rumors (Chen and Haga, 2021).

In the digital economy era, the importance of financial news on the internet, a new carrier of information transmission, is increasing. It can be seen as a supplement to information disclosure because some information contained in the financial news can give early warnings before official announcements, predict the firm performance in the future, and provide clear expectations about the inherent risks (Grewal et al., 2019). On the other hand, once an investment decision is emotionally made based on the financial news on the internet, investors may suffer wealth loss or missing investment opportunities. Studies from the behavioral finance theory reveal that the price

of any assets depends on the market participants' emotion reactions to information or news (Thaler, 2019). It needs to be emphasized that the idea that emotional reactions might influence asset prices is not a prerogative of behavioral finance theory. Psychological studies also argue that emotional responses are ubiquitous and may significantly depart from cognitive responses when facing uncertain information and news (Kahneman and Tversky, 2013). In general, biased emotions, general knowledge, and experience might account for investor reactions to information or news (Fischer, 2011), and similar studies have been mainly tested in the stock market. For example, Cervellati et al. (2014) found that the stock price and trade volume reacted to the second-hand information in Italy. Suleman (2012) analyzed the stock market's response to the political news, and their findings showed that good political news increased the stock return and decreased the volatility, while bad political news negatively impacted the abnormal returns and increased the volatility. However, little literature has focused on the bond market's reaction to financial news. Therefore, how bond investors and CRAs, two vital participants in the bond market, react to financial news on the internet should be studied.

Abundant literature has revealed that it is unique for central state-owned enterprises (SOEs). First, the chairpersons of central SOEs in China were ministers or directly reported to ministers in planned economy days. Nowadays, they still have a considerable influence at different government levels in China, especially at the highest government level. At the same time, the Communist Party of China (CPC) and central government also exert tight control over central SOEs. Most senior managers of central SOEs are members of the CPC, and they are subjected to the discipline of the CPC (Andrews-Speed, 2010). Second, central SOEs get much stronger support for business from the government, and the government is more likely to take action to avoid default when central SOEs have high risks (Huang et al., 2020a). This implies that bond investors may react differently to the news about central SOEs and other enterprises, which has barely been investigated in existing studies. This article, therefore, will extend existing literature and investigate whether the news reaction is associated with enterprise ownership.

Some scholars have discussed that there is a decline in the news effect over time. For example, Zhao and Zeng (2019) pointed out that the forecasting ability of news related to short-scale trends and long-scale trends only lasted 1 week, whereas the distinguish trends could last up to 3 weeks. Chen et al. (2021) found that the sensitivity of informed investors to the policy news would diminish over time. Thus, this article extends previous literature and attempts to investigate whether bond participants have different reactions to new news and past news.

The goal in this article is attempt to answer three research questions. The first question is whether there is a bias in the emotion reactions to bad news and good news in China's bond market and whether the emotion reactions are consistent with previous literature. The second question is whether bond investors' reactions to financial news depend on enterprise ownership, or whether bond investors have a stronger emotional reaction to financial news about central SOEs than other enterprises. The third question is whether bond investors' reactions to financial news depend on news timeless, or whether bond investors have stronger reactions to new financial news than old financial news.

We make several contributions to the literature in this study. First, investor reactions have been extensively examined (Cano et al., 2016; Narayan and Bannigidadmath, 2017; Chen et al., 2021) in the stock market, but we focused on the bond market, for which limited evidence is available. Second, prior literature mainly focused on the event news and macroeconomic level news, such as real estate event news (Chen et al., 2021), consumer price index (CPI) news (Schwert, 1981), unemployment rate (UR) news (Hakkio and Pearce, 1985), and producer price index (PPI) news (Beber and Brandt, 2010). Instead, we are interested in the financial news on the internet at each specific-enterprise level, which has a higher dimension and contains more information. Finally, we examine the heterogeneous response driven by enterprise ownership and news timeliness that bond investors differently react to central SOE news, other enterprise news, old news, new news.

The layout of the paper is organized as follows. Part 1 is the introduction. Part 2 reviews the literature. Part 3 illustrates the research design. Part 4 presents and discuss the empirical results. Part 5 offers the robust check, whereas we conclude in Part 6.

LITERATURE

Financial News and Credit Rating Agencies

Credit ratings, viewed as CRAs' reactions toward credit risk valuation, play an important role in the bond market because they correct the information asymmetry between bond issuers and financial markets (Bongaerts et al., 2012). Therefore, bond participants highly rely on credit ratings and view them as the license to the bond market (Partnoy, 2006). Abundant literature has studied the determining factors in credit ratings. For example, Altman et al. (1977), Xia (2014), and de Vries and de Haan (2016) pointed out that firms with better financial performance were more likely to receive high credit ratings. Bhandari and Golden (2021) found that CRAs had a stronger response to CEO's political ideology. Camanho et al. (2020) revealed that CRAs tended to upgrade credit ratings when facing fierce competition in the credit rating market. However, there is little literature available for how CRAs or credit ratings react to financial news on the internet. Two representative studies conducted by Tsai et al. (2010, 2016) found that the text information from the newspapers contained incremental informational content for credit risk evaluation. Therefore, they concluded that CRAs had a significant reaction to the information from the newspapers. Lu et al. (2012) revealed that news was helpful to predict future credit ratings, implying that CRAs significantly reacted to the coverages in the newspaper. In this article, we will extend the previous literature to investigate the attitudes of CRAs to the financial news on the internet.

Financial News and Bond Investors

Abundant studies have pointed out that financial news plays an important role in the financial market, and those financial news coverages are widely used and analyzed by investors (Sabherwal et al., 2011; Narayan and Bannigidadmath, 2017; Du, 2020). Narayan and Bannigidadmath (2017) examined how investors reacted to the financial news on Islamic and conventional stock markets from 2005 to 2012. They found that both positive and negative news influenced stock returns, but positive news had a relatively larger impact on stock investors. However, Du (2020) got the opposite conclusion that stock investors in Japan had a stronger reaction to negative coverages. Based on 130 oil-related words, Loughran et al. (2019) investigated how investors responded to the media news on the oil market. Their findings showed that investors often overreacted to oil news. Corbet et al. (2020) examined the relationship between news coverages and Bitcoin investors' reactions, and they found that bitcoin investors were not sensitive to CPI and GDP news but had a significant reaction to unemployment and durable goods news.

Recently, a handful of literature has focused on how the bond market reacted to financial news (Beber and Brandt, 2010; Defond and Zhang, 2014; Caporale et al., 2018). According to the contract theory, investors in the bond market were more conservative than investors in other financial markets (Kothari et al., 2010), implying that investor emotions in the bond market may be more sensitive to news coverages. Defond and Zhang (2014) tested how bond investors reacted to good and bad earning news. Their findings showed that bond investors were more likely to impound bad earning news on a timelier basis than stock investors. Beber and Brandt (2010) studied how investors responded to positive and negative macroeconomic news during economic expansions and recessions. They found that bond investors strongly responded to negative news associated with non-farm payrolls in expansions and positive news associated with inflation in recessions. Using a VAR-GARCH model, Caporale et al. (2018) investigated how bond investors reacted to macro news in the euro area from 1999 to 2014. Their findings showed that negative coverage associated with macro news positively affected bond investors in Greece, Ireland, Italy, Portugal, and Spain, and the responses became stronger in the 2008 financial crisis. However, there are several limitations in those studies. First, news used in existing literature mainly focused on the macroeconomic level, but seldom literature focused on the financial news on the internet. Second, previous literature mainly focused on the investor reactions in the secondary market rather than the investor reactions in the primary market. Therefore, this proposed article focuses on bond investors' reactions toward the financial news on the internet at each enterprise-specific level in the primary bond market.

Previous literature has discussed the difference between central SOEs and other enterprises. For example, Zhao (2006) pointed out that central SOEs play an important role and enjoy many privileges in China's economy. Lin et al. (2020) argued that central SOEs were viewed as a way to maintain social stability, without which the economy could not function properly. Ralston et al. (2006) analyzed different enterprise ownership types and concluded that SOEs contributed a significant share to the total output, particularly true for the sectors with strategic value. However, no literature associated the concept of enterprise ownership with the financial news on the internet. Considering that investor psychologies and stock market behaviors have different reactions to different news (Abreu and Mendes, 2012), we will extend previous literature and examine whether bond investors have a stronger emotion reaction to news about other enterprises than SOEs.

News timeliness has been studied extensively in previous literature. For example, Defond and Zhang (2014) examined the timeliness of the market reaction to earnings news, and they found that the timeliness of bond investors' reactions to bad news was concentrated primarily among speculative-grade bonds. Matsubara et al. (2012) and Chen et al. (2021) concluded that the news influence would diminish over time. Eachempati and Srivastava (2021) argued that investors were inclined to ignore old news that may not stir the market sentiments, but they were sensitive to new news. In this article, our study will further analyze whether bond investors have different reactions to new news and old news.

METHODOLOGY

Measure of the Reactions From Market Participants

If bond investors and CRAs have emotion reactions toward the bad (good) news, we will observe lower (higher) credit ratings or higher (lower) bond yield spreads. Thus, we employ credit ratings and bond yield spreads as the proxy variables of CRA and bond investor reactions.

The credit rating can be divided into two types. One is the bond project rating (*BR*), which can be regarded as the security of the bond itself. The other is corporate entity rating (*CR*), which can be viewed as the overall security of the bond issuer. Following Xia (2014) and Huang et al. (2020b), we employ the numerical value to quantify *BR* and *CR* as: AAA = 1, AA+ = 2, AA = 3, AA− = 4, A+ = 5, A = 6, A− = 7, BBB+ = 8, BBB = 9, BBB− = 10, BB+ = 11, BB = 12, BB− = 11, B+ = 12, B = 13, B− = 14, and lower than B− = 15.

Following Qian (2018) and Chen et al. (2021), the bond yield spread can be measured by the difference between the corporate bond yield and government bond yield, which is given as follow:

$$YieldSpread = yield_1 - yield_2 \quad (1)$$

where $_{yield_1}$ is the corporate bond yield, $_{yield_2}$ is the government bond yield with the same maturity as the corporate bond.

News Quantification

News Classification

We will classify the news into three types: positive, negative, and neutral news. First of all, three financial researchers with professional knowledge are invited to mark 100,000 news articles manually. Following Das and Chen (2007), we then apply three classifiers, namely polynomial naïve Bayes, support

vector machine (SVM), and random forest (RF), to train the marked news that is categorized manually by three experienced researchers. Finally, we employ the trained classifiers to label the unmarked news articles, and the voting scheme is used to classify the unmarked news[1].

Measure of News Sentiment

After classifying the news type, two indexes are introduced to measure the news sentiment from T_1 (6 months prior to the bond issue date) to T_0 (the bond issue date), which is denoted by:

$$BadRatio = \frac{N}{T} \tag{2-1}$$

$$GoodRatio = \frac{P}{T} \tag{2-2}$$

where T is the total number of news articles over the 6-month window, N and P are the total number of bad and good news over the 6-month window, respectively.

Next, we introduce the concept of positive/negative day to avoid false conclusions because repeated news articles could be released on different news websites on the same day. Specifically, we view the day as a positive or negative day for one specific enterprise when the enterprise is exposed to more good or bad news on one day.

Finally, to improve the precision, another two sentiment proxies are introduced to measure the proportion of bad and good news articles at the day level, which is given as follow:

$$DayBadRatio = \begin{cases} \frac{TND}{D} \ if \ D > 0 \\ 0 \ if \ D = 0 \end{cases} \tag{3-1}$$

$$DayGoodRatio = \begin{cases} \frac{TPD}{D} \ if \ D > 0 \\ 0 \ if \ D = 0 \end{cases} \tag{3-2}$$

where D is the total number of days in which the news is released on the internet for a specific enterprise, TND is the number of negative days for a specific enterprise, TPD is the number of positive days for a specific enterprise.

Econometric Models

To examine whether CRAs and bond investors have emotion reactions to financial news on the internet, we use following multiple regressions:

$$BR = \alpha + \beta_0 X + \beta_1 LnToalNews + Firm\ Control + Bond\ Control + Industry + Year + \varepsilon \tag{4}$$

$$YieldSpread = \alpha + \beta_0 X + \beta_1 LnToalNews + \beta_2 CR + Firm\ Control + Bond\ Control + Rate\ Control + Industry + Year + \varepsilon \tag{5}$$

[1]If there is an indistinguishable situation in the voting process, such as one positive vote, one neutral vote, and one negative vote, we will manually determine the final result.

TABLE 1 | Variable definition.

Variable	Definitions
Dependent variables	
CR	Corporate entity rating (AAA = 1, AA+ = 2, AA = 3, AA− = 4, A+ = 5, A = 6, A− = 7, BBB+ = 8, BBB = 9, BBB− = 10, BB+ = 11, BB = 12, BB− = 11, B+ = 12, B = 13, B− = 14, and lower than B− = 15)
BR	Bond project rating (AAA = 1, AA+ = 2, AA = 3, AA− = 4, A+ = 5, A = 6, A− = 7, BBB+ = 8, BBB = 9, BBB− = 10, BB+ = 11, BB = 12, BB− = 11, B+ = 12, B = 13, B− = 14, and lower than B− = 15)
YieldSpread	The difference between the government bond yield and the municipal corporate bond yield, with the same maturity
Gua	Whether the bond issuer seeks the credit guarantee or not?
Firm Control	
LnAsset	The natural logarithm of MC's total assets with one lagged year
Liability	The asset-liability ratio with one lagged year
Current	The current ratio with one lagged year
ROE	The return on equity with one lagged year
TurnOver	The total assets turnover with one lagged year
SOE	The enterprise ownership (central SOEs and other enterprises)
Bond Control	
BondSize	Bond size
Maturity	Bond maturity
Collateral	Can the bond be collateral (Yes: 1; No: 0)?
BondType	Bond Type[1] (including the enterprise bond, corporate bond, and medium-term note)
Rate Control	
GovYield	Government bond yield with the same maturity as the corporate bond
OneYearRate	1-year government bond yield
News variable	
BadRatio	See Part News Quantification
GoodRatio	See Part News Quantification
DayBadRatio	See Part News Quantification
DayGoodRatio	See Part News Quantification
LnTotalNews	The natural logarithm of total news articles over 6 months
Other control variable	
Industry	Industry
Year	Year

[1]*Enterprise bonds approved by the National Development and Reform Commission (NDRC) are traded in the exchange and interbank markets. Corporate bonds approved by the China Securities Regulatory Commission (CSRC) are traded in the exchange market.*

where X is the variable related to news sentiment indexes, including *BadRatio*, *GoodRatio*, *DayBadRatio*, and *DayGoodRatio*. $_{LnToalNews}$ is the nature logarithm of the total number of news articles. *CR* is the corporate entity rating. *Firm Control* includes the natural logarithm of total assets, asset-liability ratio, total assets turnover, current ratio, return on equity, and enterprise ownership. *Bond Control* includes bond size, bond maturity, collateral clause, and bond type. *Rate Control* includes two yields: 1-year government bond yield and government bond yield with the same maturity as the corporate bond. Year and industry fixed effects control specific time and industry factors. **Table 1** describes these variables in detail.

Data Statistics

We collect the data from several sources, ranging from 2010 to 2020. Data related to bond and firm characteristics are downloaded from Wind[2]. 275,470 news articles are collected from China's mainstream financial news websites using web-crawler technology. Deleting some observations with missing information, we finally select 2,510 bonds (central SOEs issue 690 bonds and other enterprises issue 1,820 bonds) issued by listed enterprises in China, excluding financial bonds, convertible bonds, notes, and asset-backed securities. **Tables 2**, **3** report the statistical results based on Stata 15.

Panel A of **Table 2** reports the statistics about the whole news sample. The mean of *TotalNews* is 91.37, but the median is 19. *BadRatio* (mean of 0.181, median of 0.146) and *DayBadRatio* (mean of 0.209, median of 0.177) are relatively smaller than *GoodRatio* (mean of 0.683, median of 0.714) and *DayGoodRatio* (mean of 0.679, median of 0.700), indicating that more positive financial news is released on the internet.

Panel B of **Table 2** provides the statistics about central SOEs. We find that central SOEs have relatively higher bond project ratings (mean of 1.268, median of 1), higher corporate entity ratings (mean of 1.362, median of 1), and lower bond yield spreads (mean of 224, median of 193). In addition, we also find that central SOEs have more positive news than negative news, which is consistent with Panel A.

Panel C of **Table 2** presents the statistics about other enterprises. Consistent with Panel A and Panel B, other enterprises are also exposed to more good news. On the other hand, other enterprises have relatively lower bond project ratings (mean of 1.973, median of 2), lower corporate entity ratings

[2]https://www.wind.com.cn/

TABLE 2 | Descriptive statistics of the internet news about China's listed enterprises from 2010 to 2020.

	Panel A: All sample			Panel B: Central SOEs			Panel C: Other enterprises		
Variable	***N***	**Mean**	**p50**	***N***	**Mean**	**p50**	***N***	**Mean**	**p50**
BR	2394	1.772	1	682	1.268	1	1712	1.973	2
CR	2510	2.022	2	690	1.362	1	1820	2.271	2
YieldSpread	2510	224.0	193	690	138.8	132	1820	256.2	239
TotalNews	2510	91.37	19	690	133.9	45	1820	75.25	16
LnTotalNews	2510	3.146	2.944	690	3.599	3.806	1820	2.974	2.773
BadRatio	2510	0.181	0.146	690	0.192	0.167	1820	0.177	0.143
GoodRatio	2510	0.683	0.714	690	0.664	0.691	1820	0.691	0.721
DayBadRatio	2510	0.209	0.177	690	0.222	0.200	1820	0.204	0.167
DayGoodRatio	2510	0.679	0.700	690	0.666	0.667	1820	0.684	0.714

TABLE 3 | Pearson correlation coefficients.

	YieldSpread	**CR**	**LnTotalNews**	**BadRatio**	**GoodRatio**	**DayBadRatio**	**DayGoodRatio**
YieldSpread	1						
CR	0.575	1					
LnTotalNews	−0.184	−0.300	1				
BadRatio	0.0103	0.0473	0.0340	1			
GoodRatio	−0.0009	−0.0051	−0.0168	−0.753	1		
DayBadRatio	0.0032	0.0273	0.0971	0.960	−0.736	1	
DayGoodRatio	−0.0162	−0.0204	0.0236	−0.747	0.949	−0.775	
LnAsset	−0.360	−0.749	0.373	−0.0613	0.0140	−0.0284	
Liability	0.0986	−0.0295	0.0603	−0.0632	0.0598	−0.0510	
ROE	0.0069	−0.115	−0.0024	−0.139	0.105	−0.141	
Current	0.191	0.268	−0.102	−0.041	0.0405	−0.0396	
TurnOver	−0.0147	−0.0241	0.0276	0.0404	−0.0353	0.0393	
Gua	0.150	0.358	−0.0574	0.132	−0.112	0.114	
	DayGoodRatio	**LnAsset**	**Liability**	**ROE**	**Current**	**TurnOver**	**Gua**
DayGoodRatio	1						
LnAsset	0.0258	1					
Liability	0.0515	0.216	1				
ROE	0.108	0.106	−0.266	1			
Current	0.0310	−0.286	−0.0302	0.0989	1		
TurnOver	−0.0252	0.0327	−0.0582	−0.0711	−0.251	1	
Gua	−0.102	−0.305	0.0228	−0.128	0.108	0.0404	1

(mean of 2.271, median of 2), and higher bond yield spreads (mean of 256.2, median of 239).

Table 3 reports the results of the Pearson correlation coefficients between the main variables. We can see that the correlation coefficient between *LnTotalNews* and *CR* is −0.3, indicating that bonds with higher credit ratings are exposed to more financial news. In addition, we find the negative/positive correlation between the bond yield spread and bad/good news, implying that bond investors may have negative/positive reactions to bad/good news.

RESULTS AND DISCUSSION

Preliminary Findings

Firstly, we study whether CRAs and bond investors have different reactions to financial news on the internet. The empirical results are shown in **Table 4**.

Columns (1)–(4) of **Table 4** report the results of CRA reactions. Results show that the coefficients of *BadRatio*, *GoodRatio*, *DayBadRatio*, and *DayGoodRatio* are insignificant, indicating that financial news on the internet does not influence CRAs. This is inconsistent with Tsai et al. (2010), who found that news impacted CRAs.

The results may be interpreted as follows. First, CRAs extract the credit rating fees from bond issuers, whereas bond issuers can choose any CRAs they prefer. In this case, it is not strange that CRAs have motivations to underestimate financial news on the internet to please bond issuers and maintain the market share (Xia, 2014). Second, the financial news from the internet contains high noise and potential biases, which could mislead the market participants (Chen and Haga, 2021). Therefore, CRAs keep calm and tend to not adjust credit ratings before they see official announcements.

Columns (5)–(8) of **Table 4** report the results of bond investor reactions. Results indicate that bond investors have significant reactions to both bad and good news. On average, 1% increase in bad news leads to an increase in the bond yield spread by 0.26 BP (column 5), while 1% increase in good news drives the bond yield spread down by 0.22 BP (column 6), indicating that the bad news has a more significant influence on bond investors. In addition, the coefficient of *DayBadRatio* (26.249, with a *t*-value of 3.01) and *DayGoodRatio* (−23.467, with a *t*-value of −3.31) in columns (7)–(8) show a similar conclusion that bond investors have an asymmetric reaction to bad and good news.

Overall, the empirical results indicate that bond investors have a stronger reaction to bad news than good news, which is similar to Du (2020) who found that negative news had a stronger influence on Japan's stock market. Our results are also consistent with studies in psychology that negative information has a stronger impact on impressions than positive information (Baumeister et al., 2001). A number of explanations account for the asymmetric reaction have been put forward. The negative news bias may result from evolution as attention to negative information makes survival easier, and increases the possibility of genetic inheritance (Baumeister et al., 2001). The negative news bias also could be explained by loss aversion that investors care more about a loss of utility than a gain of equal magnitude (Kahneman and Tversky, 2013).

TABLE 4 | Result about how CRAs and bond investors react toward financial news on the internet.

Variables	(1) BR	(2) BR	(3) BR	(4) BR	(5) YieldSpread	(6) YieldSpread	(7) YieldSpread	(8) YieldSpread
BadRatio	−0.013				26.315***			
	(−0.19)				(2.86)			
GoodRatio		0.003				−22.166***		
		(0.05)				(−3.02)		
DayBadRatio			0.015				26.249***	
			(0.23)				(3.01)	
DayGoodRatio				−0.012				−23.467***
				(−0.23)				(−3.31)
LnTotalNews	0.016	0.016	0.016	0.016	−2.655*	−2.781*	−2.886*	−2.642*
	(1.59)	(1.58)	(1.55)	(1.57)	(−1.68)	(−1.75)	(−1.83)	(−1.68)
CR					56.622***	56.433***	56.410***	56.380***
					(13.49)	(13.4)	(13.43)	(13.40)
Firm Control	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,394	2,394	2,394	2,394	2,510	2,510	2,510	2,510
Adj R^2	0.627	0.627	0.627	0.627	0.56	0.561	0.561	0.561

*This table includes regressions to test how CRAs and bond investors react to financial news on the internet based on Stata 15. Standardized betas are reported and p-values are presented in parentheses. Symbols of ***, **, and * represent significance at the 1, 5, and 10% level, respectively.*

Further Analysis of Enterprise Ownership

It is not amazing that central SOEs enjoy a tremendous news advantage in China. For example, central SOEs are more likely to get media attention and be exposed to positive news. Therefore, we divide the news into two groups: central SOE news and other enterprise news. Then we perform the interaction analysis to study the heterogeneous reactions under different enterprise ownerships. Particularly, we add four interaction terms, *Dum_SOE* × *BadRatio*, *Dum_SOE* × *GoodRatio*, *Dum_SOE* × *DayBadRatio*, and *Dum_SOE* × *DayGoodRatio*, to capture the differential reactions. *Dum_SOE* is a dummy variable, which is equal to 1 when the news is about central SOEs, but 0 otherwise. **Table 5** shows the empirical results.

Column (1) shows that the coefficient of *Dum_SOE* × *BadRatio* is 30.919 with a *t*-value of 1.75, indicating that bond investors have stronger reactions to negative news about other enterprises.

Column (2) shows that the coefficient of *Dum_SOE* × *GoodRatio* is −32.737 with a significance at 10% level, indicating that bond investors strongly react to positive news about other enterprises. Columns (3)–(4) show a similar result that the coefficients of *Dum_SOE* × *DayBadRatio* and *Dum_SOE* × *DayGoodRatio* are 40.429 and −34.796, respectively, both significant at 5% level.

Next, we divide the sample into two subsamples: bonds issued by central SOEs and other enterprises. Then we conduct the regression analysis based on the two subsamples and compare the results with the interaction analysis. The empirical results are shown in **Table 6**.

In columns (1), (3), (5), and (7), the bonds are issued by central SOEs, and results show that the coefficients of *GoodRatio*, *BadRatio*, *DayGoodRatio*, and *DayBadRatio* are insignificant, indicating that bond investors are not influenced by the news about central SOEs.

In columns (2), (4), (6), and (8), the bonds are issued by other enterprises, and the results show that the coefficients of *GoodRatio*, *BadRatio*, *DayGoodRatio*, and *DayBadRatio* are all significant at 5% level, indicating that bond investors are more sensitive to the news about other enterprises than SOEs. This is also consistent with the results in **Table 5**.

Overall, our results confirm that there is heterogeneity in the psychological reaction where bond investors do not react to news about SOEs but to the news about other enterprises. Two reasons could explain our result. First, investors have adjusted their sensitivity to the news about central SOEs as too much news about central SOEs is released on the internet every day. Second, central SOEs could get strong support and implicit guarantees from governments (Huang et al., 2020a), implying that bonds issued by central SOEs are secure. Thus, market participants are not sensitive to the news about central SOEs.

TABLE 5 | Interaction analysis.

Variables	(1) YieldSpread	(2) YieldSpread	(3) YieldSpread	(4) YieldSpread
BadRatio	2.593			
	(0.18)			
Dum_SOE × BadRatio	30.919*			
	(1.75)			
GoodRatio		3.259		
		(0.29)		
Dum_SOE × GoodRatio		−32.737**		
		(−2.34)		
DayBadRatio			−5.498	
			(−0.41)	
Dum_SOE × DayBadRatio			40.429**	
			(2.42)	
DayGoodRatio				3.564
				(0.33)
Dum_SOE × DayGoodRatio				−34.796**
				(−2.56)
LnTotalNews	−2.580	−2.534	−2.625*	−2.414
	(−1.62)	(−1.59)	(−1.65)	(−1.52)
CR	56.358***	56.359***	56.282***	56.222***
	(13.40)	(13.41)	(13.42)	(13.38)
Firm Control	Yes	Yes	Yes	Yes
Bond Control	Yes	Yes	Yes	Yes
Rate Control	Yes	Yes	Yes	Yes
Industry	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes
Observations	2,510	2,510	2,510	2,510
Adj R^2	0.561	0.561	0.561	0.562

*This table includes regressions to test the interaction effects based on Stata 15. Dum_SOE is a dummy variable which is equal to 1 if the news is related to central SOEs, but 0 otherwise. Dum_SOE × BadRatio, Dum_SOE × GoodRatio, Dum_SOE × DayBadRatio, and Dum_SOE × DayGoodRatio capture the differential impact of enterprise ownership about the bad news and good news. Standardized betas are reported and p-values are presented in parentheses. Symbols of ***, **, and * represent significance at the 1, 5, and 10% level, respectively.*

Further Analysis of Old News and New News

Some news can affect the bond investors for only a long period, whereas some news has a short-term impact on the bond investors. To examine whether both new news and old news have different influences on bond investors, we divide the 6-month window into two 3-month windows, which is shown in **Figure 1**.

In **Figure 1**, *T0* is the date when we issue the bond. We view the news in the first 3-month window as the past or old news, while news in the second 3-month window is viewed as the latest news or new news. Next, we construct the four new proxies to capture the news sentiment based on the two sub-windows, which is given as follow:

$$DayBadRatio_i = \begin{cases} \frac{TND_i}{D_i} \ if \ D_i > 0 \\ 0 \ if \ D_i = 0 \end{cases} \quad (6\text{-}1)$$

TABLE 6 | Investor reactions based on central SOEs and other enterprises.

Variables	(1) YieldSpread	(2) YieldSpread	(3) YieldSpread	(4) YieldSpread	(5) YieldSpread	(6) YieldSpread	(7) YieldSpread	(8) YieldSpread
BadRatio	−2.972	26.209**						
	(−0.26)	(2.13)						
GoodRatio			5.378	−18.736**				
			(0.55)	(−1.97)				
DayBadRatio					−3.057	23.947**		
					(−0.28)	(2.09)		
DayGoodRatio							7.758	−20.111**
							(0.85)	(−2.17)
LnTotalNews	−2.150*	−0.192	−2.086*	−0.162	−2.133*	−0.309	−2.094*	−0.034
	(−1.78)	(−0.08)	(−1.71)	(−0.07)	(−1.75)	(−0.13)	(−1.74)	(−0.01)
CR	46.461***	74.632***	46.474***	74.623***	46.485***	74.594***	46.354***	74.552***
	(8.62)	(13.92)	(8.74)	(13.91)	(8.69)	(13.93)	(8.68)	(13.90)
Bond Control	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm Control	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Rate Control	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	690	1,820	690	1,820	690	1,820	690	1,820
Adj R^2	0.612	0.403	0.613	0.403	0.612	0.403	0.613	0.403

*This table includes regressions to test how bond investors react to financial news on the internet based on Stata 15. The bonds in columns (1), (3), (5), and (7) are issued by central SOEs, whereas the bonds in columns (2), (4), (6), and (8) are issued by other enterprises. Standardized betas are reported and p-values are presented in parentheses. Symbols of ***, **, and * represent significance at the 1, 5, and 10% level, respectively.*

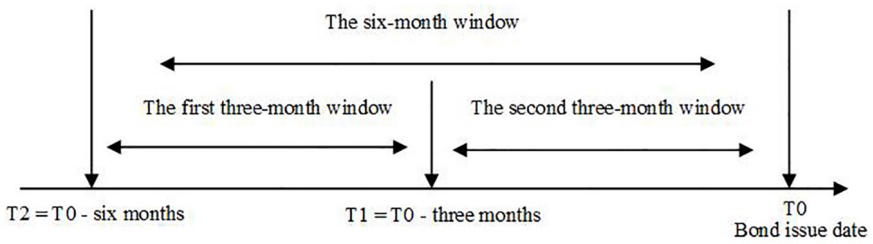


FIGURE 1 | Dividing the 6-month window into two 3-month windows.

$$DayGoodRatio_i = \begin{cases} \frac{TPD_i}{D_i} \ if \ D_i > 0 \\ 0 \ if \ D_i = 0 \end{cases} \quad (6\text{-}2)$$

$$CDayBadRatio_i = \begin{cases} \frac{TND_i}{D} \ if \ D > 0 \\ 0 \ if \ D = 0 \end{cases} \quad (6\text{-}3)$$

$$CDayGoodRatio_i = \begin{cases} \frac{TPD_i}{D} \ if \ D > 0 \\ 0 \ if \ D = 0 \end{cases} \quad (6\text{-}4)$$

where i is equal to 1 if the news is released in the first 3-month window, but 2 in the second 3-month window. $_{TND_i}$ is the total number of negative days in window i.$_{TPD_i}$ is the total number of positive days in window i. $_{D_i}$ is the total days in which the financial news is released on the internet for one specific enterprise in window i. D is the total days in which the financial news is released on the internet for one specific enterprise in the 6-month window.

Finally, we examine whether bond investors react to financial news based on the news in two sub-windows. The empirical results are shown in **Table 7**.

Columns (1)–(2) show that the coefficients of $DayBadRatio_1$ and $DayGoodRatio_1$ are both significant at 5% level, indicating bond investors react to bad and good news in the first 3-month window. Columns (3)–(4) show a similar result that news in the second 3-month window also exhibits a significant influence on bond investors. In column (5), the coefficient of $CBadDayRatio_2$ (29.251, with a t-value of 2.23) is large than the coefficient of $CBadDayRatio_1$ (23.641, with a t-value of 2.11), indicating that the latest news rather than past news has a stronger influence on bond investors. A similar pattern is observed about the coefficients of $CGoodDayRatio_2$ (−24.860, with a t-value of −2.92) and $CGoodDayRatio_1$ (−22.341, with a t-value of −2.74) in column (6).

Overall, our results confirm the asymmetric response driven by news timeliness that bond investors are more

TABLE 7 | Investor reaction to old news and new news.

Variables	(1) YieldSpread	(2) YieldSpread	(3) YieldSpread	(4) YieldSpread	(5) YieldSpread	(6) YieldSpread
$DayBadRatio_1$	15.978**					
	(2.16)					
$DayGoodRatio_1$		−12.017**				
		(−2.19)				
$LnTotalNews_1$	−2.930*	−1.919				
	(−1.86)	(−1.22)				
$DayBadRatio_2$			15.738**			
			(2.14)			
$DayGoodRatio_2$				−13.698**		
				(−2.53)		
$LnTotalNews_2$			−2.640	−1.066		
			(−1.64)	(−0.65)		
$CBadDayRatio_1$					23.641**	
					(2.11)	
$CBadDayRatio_2$					29.521**	
					(2.23)	
$CGoodDayRatio_1$						−22.341***
						(−2.74)
$CGoodDayRatio_2$						−24.860***
						(−2.92)
LnTotalNews					−2.905*	−2.616*
					(−1.84)	(−1.65)
CR	56.513***	56.502***	56.396***	56.271***	56.392***	56.377***
	(13.44)	(13.41)	(13.41)	(13.40)	(13.42)	(13.40)
Firm Control	Yes	Yes	Yes	Yes	Yes	Yes
Bond Control	Yes	Yes	Yes	Yes	Yes	Yes
Rate Control	Yes	Yes	Yes	Yes	Yes	Yes
Industry	Yes	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,510	2,510	2,510	2,510	2,510	2,510
Adj R^2	0.560	0.560	0.560	0.560	0.561	0.561

*This table includes regressions to test how bond investors react to old news and new news on the internet based on Stata 15. Standardized betas are reported and p-values are presented in parentheses. Symbols of ***, **, and * represent significance at the 1, 5, and 10% level, respectively.*

sensitive to new news than old news, which is within our expectations. Two explanations account for this phenomenon. First, old news contains gossip or uncertain information, which may fluctuate investor emotions in the past (Chen and Haga, 2021). However, as time goes by, the information in old news may have no fundamental values, suggesting that the influence of old news on bond investors will decline over time. Second, new news contains more risk and uncertainty, leading an emotional decision (Stracca, 2004).

ROBUST CHECK

Alternative Measure of Bond Yield Spread

Previous section in our paper calculates the bond yield spread based on the nominal interest rate when the bond is issued. In the first robust check, we use the average closing price in the secondary bond market to estimate the bond yield spread. Then we again perform the regressions to test investor reactions to financial news on the internet. **Table 8** presents the results.

In columns (1)–(4), the bonds are issued by other enterprises. We observe that the coefficients of *BadRatio*, *GoodRatio*, *DayBadRatio*, and *DayGoodRatio* are significant at 1% level. In addition, the absolute value of the coefficients of *BadRatio* and *DayBadRatio* are relatively larger than the coefficients of *GoodRatio* and *DayGoodRatio*. Conversely, the bonds in columns (5)–(6) are issued by central SOEs, and we find that the coefficients of news sentiment variables are insignificant. Overall, the result confirms our previous conclusions.

Alternative Measure of News Sentiment

The method in our empirical study to classify news may misclassify news type. In this subsection, we follow

TABLE 8 | Robust check 1.

Variables	(1) YieldSpread	(2) YieldSpread	(3) YieldSpread	(4) YieldSpread	(5) YieldSpread	(6) YieldSpread	(7) YieldSpread	(8) YieldSpread
BadRatio	36.394***				3.177			
	(2.74)				(0.23)			
GoodRatio		−27.767***				0.862		
		(−2.72)				(0.08)		
DayBadRatio			35.059***				1.720	
			(2.89)				(0.13)	
DayGoodRatio				−29.443***				3.738
				(−3.00)				(0.36)
LnTotalNews	−2.142	−2.006	−2.324	−1.832	−1.880	−1.803	−1.862	−1.766
	(−0.84)	(−0.79)	(−0.91)	(−0.72)	(−1.38)	(−1.31)	(−1.35)	(−1.31)
CR	61.533***	61.416***	61.436***	61.233***	45.326***	45.174***	45.248***	45.068***
	(11.00)	(10.97)	(11.01)	(10.94)	(7.77)	(7.86)	(7.82)	(7.82)
Firm Control	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Bond Control	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Rate Control	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,629	1,629	1,629	1,629	653	653	653	653
Adj R^2	0.496	0.496	0.496	0.496	0.571	0.571	0.571	0.571

*This table includes regressions to test how CRAs and bond investors react to financial news on the internet based on Stata 15. Standardized betas are reported and p-values are presented in parentheses. Symbols of ***, **, and * represent significance at the 1, 5, and 10% level, respectively.*

TABLE 9 | Robust check 2.

Variables	(1) BR	(2) YieldSpread	(3) YieldSpread	(4) YieldSpread	(5) YieldSpread	(6) YieldSpread
div	0.007	13.040***	−1.693	17.258***		
	(0.33)	(4.83)	(−0.46)	(5.49)		
div1					6.909***	
					(2.99)	
div2						7.666***
						(3.32)
LnTotalNews	−0.001	−2.388	−2.221*	−1.106	−1.382	−1.736
	(−0.08)	(−1.52)	(−1.91)	(−0.48)	(−0.85)	(−1.10)
CR		56.124***	46.439***	61.105***	56.184***	56.235***
		(13.33)	(8.75)	(11.66)	(13.37)	(13.35)
Firm Control	Yes	Yes	Yes	Yes	Yes	Yes
Bond Control	No	Yes	Yes	Yes	Yes	Yes
Rate Control	No	Yes	Yes	Yes	Yes	Yes
Industry	Yes	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,394	2,510	690	1,820	2,510	2,510
Adj R^2	0.624	0.563	0.612	0.491	0.561	0.561

*This table includes regressions to test how CRAs and bond investors react to financial news on the internet based on Stata 15. Standardized betas are reported and p-values are presented in parentheses. Symbols of ***, **, and * represent significance at the 1, 5, and 10% level, respectively.*

Chen et al. (2021) and measure the news sentiment as follow:

$$div = \log \frac{1+N}{1+P} \tag{7}$$

where N and P are the total number of negative and positive words for a specific enterprise.

Then we again perform the regression based new news sentiment variable. **Table 9** presents the results. In columns (1)–(5), we calculate *div* based on the 6-month window. In column (6), *div1* and *div2* are calculated based on the first and second 3-month window, respectively.

Column (1) shows that the coefficient of *div* is insignificant, indicating that CRAs have no reactions to financial news on

the internet. In columns (2)–(4), bonds are issued by all listed enterprises, central SOEs, and other enterprises, respectively. It is observed that the coefficients of *div* are negative in column (2) and column (4), both significant at 1% level, while the coefficient of *div* in column (3) is insignificant, indicating that bond investors are sensitive to the news about other enterprises but not to the news about central SOEs. Columns (5)–(6) show that the coefficients of *div1* and *div2* are 6.909 and 7.666, both significant at 1% level, suggesting that both old and new news influence bond investors, but bond investors have strong reactions to new news.

Endogenous Test

We also employ the instrumental variables (IVs) and conduct a two-stage least square (2SLS) estimation to avoid the endogenous problem. We introduce two variables that are theoretically non-related to the bond yield spread and bond rating but associated with *DayBadRatio*, *DayGoodRatio*, *BadRatio*, and *GoodRatio* as our IVs. These IVs are *IV_BadRatio*, *IV_GoodRatioIV*, *IV_BadRatio*, and *IV_GoodRatio*, which are the average of *DayBadRatio*, *DayGoodRatio*, *BadRatio*, and *GoodRatio* of other bond issuers with the same year, industry, corporate entity rating. A larger IV captures a higher likelihood of exposure to the medium, reflecting that bond issuers with similar performance and industry could get similar media attention. Next, we employ a first-stage probit model, as specified in Model (8), to estimate the predicted values of *DayBadRatio*, *DayGoodRatio*, *BadRatio*, and *GoodRatio*, respectively, which are then used as an instrument in the second-stage least squares estimation.

$$X = \alpha + \beta_0 IVX + \beta_1 LnToalNews + Firm\ Control + Industry + Year + \varepsilon \quad (8)$$

where X is *DayBadRatio*, *DayGoodRatio*, *BadRatio*, or *GoodRatio*, *IVX* is our IV candidates, which are *IV_BadRatio*, *IV_GoodRatioIV*, *IV_BadRatio*, or *IV_GoodRatio*.

TABLE 10 | Robust check 3.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Variables	**Stage 1 DayBadRatio**	**Stage 2 YieldSpread**	**Stage 1 DayGoodRatio**	**Stage 2 YieldSpread**	**Stage 1 BadRatio**	**Stage 2 YieldSpread**	**Stage 1 GoodRatio**	**Stage 2 YieldSpread**
Hat_DayBadRatio		68.436***						
		(3.51)						
Hat_DayGoodRatio				−43.243***				
				(−2.75)				
Hat_BadRatio						67.202***		
						(3.26)		
Hat_GoodRatio								−38.930**
								(−2.38)
IV_DayBadRatio	0.960***							
	(25.90)							
IV_DayGoodRatio			0.975***					
			(31.78)					
IV_BadRatio					0.967***			
					(24.58)			
IN_GoodRatio							0.971***	
							(29.60)	
LnTotalNews	0.014***	−3.660**	−0.007	−2.844*	0.011***	−3.282**	−0.013***	−3.084*
	(3.94)	(−2.20)	(−1.49)	(−1.74)	(3.13)	(−1.99)	(−2.79)	(−1.87)
CR	0.005	55.721***	−0.016**	55.904***	0.002	55.920***	−0.015**	56.004***
	(0.85)	(18.78)	(−2.25)	(18.82)	(0.29)	(18.86)	(−2.16)	(18.84)
Firm Control	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Bond Control	No	Yes	No	Yes	No	Yes	No	Yes
Rate Control	No	Yes	No	Yes	No	Yes	No	Yes
Industry	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,554	2,510	2,554	2,510	2,554	2,510	2,554	2,510
Adj R^2	0.259	0.561	0.252	0.560	0.262	0.561	0.256	0.560

*This table includes 2SLS regressions to test bond investors' reactions to financial news on the internet based on Stata 15. Column (1), column (3), column (5), and column (7) show the first-stage results, whereas column (2), column (4), column (6) and column (8) show the second-stage results. IV_DayBadRatio, IV_DayGoodRatio, IV_BadRatio, and IN_GoodRatio are four IVs. Hat_DayBadRatio, Hat_DayGoodRatio, Hat_BadRatio, and Hat_GoodRatio are four predicted values from the first-stage regressions. Standardized betas are reported and p-values are presented in parentheses. Symbols of ***, **, and * represent significance at the 1, 5, and 10% level, respectively.*

Table 10 presents the 2SLS regression results. Columns (1), (3), (5), and (7) report the first-stage regression results, and we find the significant correlation between the four news indexes (*DayBadRatio*, *DayGoodRatio*, *BadRatio*, and *GoodRatio*) and our IV candidates (*IV_BadRatio*, *IV_GoodRatioIV*, *IV_BadRatio*, and *IV_GoodRatio*), thereby proving our instrumental variables' relevancy. Columns (2), (4), (6), and (8) present the second-stage regression results, where *DayBadRatio, DayGoodRatio, BadRatio, and GoodRatio* in the baseline regression are replaced with the predicted value, *Hat_DayBadRatio, Hat_DayGoodRatio, Hat_BadRatio, and Hat_GoodRatio*, from the first-stage regression. Column (2) and column (6) show that the coefficient of *Hat_DayBadRatio* (coefficient: 68.436; *t*-value: 3.51) and *Hat_BadRatio* (coefficient: 67.202; *t*-value: 3.26) remains positive and significant at 1% level. Column (4) and column (8) show a similar results that the coefficients of *Hat_DayGoodRatio* (coefficient: −43.243; *t*-value: 2.75) and *Hat_GoodRatio* (coefficient: −38.93; *t*-value: −2.38) are significant. Overall, these findings confirm that bad (good) news can significantly affect the bond investor reactions, consistent with our previous results.

CONCLUSION

This article attempts to test the psychological reactions of CRAs and bond investors to the financial news on the internet in China's bond market. As we know, much negative and credible news could be observed on the internet before the bond defaults, but CRAs do not downgrade the bond issuer. On the other hand, the bond price may fall when some negative news about the bond issuer is released on the internet. These phenomena indicate that CRAs and bond investors may have different psychological reactions to financial news. Therefore, we attempt to examine how CRAs and bond investors react to financial news on the internet.

Our study has several interesting findings. First, it demonstrates that CRAs and bond investors react differently to financial news on the internet. Specifically, CRAs display a tendency to underreact to financial news, consistent with the phenomenon that bonds exposed to bad financial news could still receive high credit ratings. However, there exists an asymmetry in psychological reactions where bond investors react stronger to bad news than to good news, consistent with the studies in behavioral finance and psychology. Second, bond investors' reactions to financial news with heterogeneity that is reflected in the enterprise ownership and news timeliness. Specifically, investors react to news about central enterprises but not news about other enterprises. In addition, investors are more sensitive to new news than old news. Overall, the study's practical contribution helps regulators obtain bond participants' reactions to financial news and formulate corresponding regulatory rules quickly. Our research also provides a reference for CRAs and investors to know psychological reactions so that they avoid overreacting or underreacting to the news.

This study has several limitations. One limitation is that we conduct our empirical study based on the listed enterprises in China, which limits its generalizability. Bonds issued by unlisted enterprises are important parts of the bond market. Usually, unlisted enterprises receive little media attention and relatively low credit ratings, and Defond and Zhang (2014) argued that bond investors were more sensitive to junk or low-rated bonds. Therefore, bond investors may have heterogeneous reactions to news about listed and unlisted enterprises. Another limitation is that we do not classify the types of news in detail. For example, Shepperd and McNulty (2002) revealed that investors reacted differently to the expected and unexpected news. However, we do not distinguish between expected and unexpected news in this study. In the future, we will study how CRAs and bond investors react to expected and unexpected financial news on the internet based on the listed and unlisted enterprises.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

WZ designed, drafted, and revised the manuscript. JW collected the data, designed the study, and revised the manuscript. MT designed, reviewed, and revised the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the National Natural Science Foundation of China (NSFC) (72071160, 62072379, and 714 73201), Fundamental Research Funds for the Central Universities (kjcx20210103 and JBK2103016), Financial Intelligence and Financial Engineering Key Lab of Sichuan Province, Chengdu SWUFE Jiaozi Institute of Fintech Innovation Co., Ltd. (cgzh20210204), Research Program of Science and Technology at Universities of Inner Mongolia Autonomous Region (2021GG0164), and Financial Innovation Center of the Southwestern University of Finance and Economics.

REFERENCES

Abreu, M., and Mendes, V. (2012). Information, overconfidence and trading: do the sources of information matter? *J. Econ. Psychol.* 33, 868–881. doi: 10.1016/j.joep.2012.04.003

Altman, E. I., Haldeman, R. G., and Narayanan, P. (1977). ZETATM analysis a new model to identify bankruptcy risk of corporations. *J. Bank. Finance* 1, 29–54. doi: 10.1016/0378-4266(77)90017-6

Andrews-Speed, P. (2010). *The Institutions of Energy Governance in China. Institut Français des Relations Internationales*. Available online at: https://www.ifri.org/sites/default/files/atoms/files/noteandrewsspeedenergychina_1.pdf (accessed February 15, 2022).

Baumeister, R. F., Bratslavsky, E., Finkenauer, C., and Vohs, K. D. (2001). Bad is str onger than good. *Rev. Gen. Psychol.* 5, 323–370. doi: 10.1037/1089-2680.5.4.323

Beber, A., and Brandt, M. W. (2010). When it cannot get better or worse: the asymmetric impact of good and bad news on bond returns in

expansions and recessions. *Rev. Finance* 14, 119–155. doi: 10.1093/rof/rfp006

Bhandari, A., and Golden, J. (2021). CEO political preference and credit ratings. *J. Corp. Finance* 68:101909. doi: 10.1016/j.jcorpfin.2021.101909

Bongaerts, D., Cremers, K. J. M., and Goetzmann, W. N. (2012). Tiebreaker: certification and multiple credit ratings. *J. Finance* 67, 113–152. doi: 10.1111/j.1540-6261.2011.01709.x

Camanho, N., Deb, P., and Liu, Z. (2020). Credit rating and competition. *Int. J. Finance Econ.* in press doi: 10.1002/ijfe.2303

Cano, C., Jareño, F., and Tolentino, M. (2016). Investor behavior and flow-through capability in the US stock market. *Front. Psychol.* 7:668. doi: 10.3389/fpsyg.2016.00668

Caporale, G. M., Spagnolo, F., and Spagnolo, N. (2018). Macro news and bond yield spreads in the euro area. *Eur. J. Finance* 24, 114–134. doi: 10.1080/1351847x.2017.1285797

Cervellati, E. M., Ferretti, R., and Pattitoni, P. (2014). Market reaction to second-hand news: inside the attention-grabbing hypothesis. *Appl. Econ.* 46, 1108–1121. doi: 10.1080/00036846.2013.866206

Chen, F., Huang, J., Sun, Z., and Tong, Y. (2020). Why do firms issue guaranteed bonds. *J. Bank. Finance* 119:105396. doi: 10.1016/j.jbankfin.2018.08.002

Chen, S. T., and Haga, K. Y. A. (2021). Using E-GARCH to analyze the impact of investor sentiment on stock returns near stock market crashes. *Front. Psychol.* 12:664849. doi: 10.3389/fpsyg.2021.664849

Chen, Y., Xie, Z., Zhang, W., Xing, R., and Li, Q. (2021). Quantifying the effect of real estate news on Chinese stock movements. *Emerg. Markets Finance Trade* 57, 4185–4210. doi: 10.1080/1540496x.2019.1695596

Corbet, S., Larkin, C., Lucey, B. M., Meegan, A., and Yarovaya, L. (2020). The impact of macroeconomic news on Bitcoin returns. *Eur. J. Finance* 26, 1396–1416. doi: 10.1080/1351847x.2020.1737168

Das, S. R., and Chen, M. Y. (2007). Yahoo! for Amazon: sentiment extraction from small talk on the web. *Manag. Sci.* 53, 1375–1388. doi: 10.1287/mnsc.1070.0704

de Vries, T., and de Haan, J. (2016). Credit ratings and bond spreads of the GIIPS. *Appl. Econ. Lett.* 23, 107–111. doi: 10.1080/13504851.2015.1054063

Defond, M. L., and Zhang, J. (2014). The timeliness of the bond market reaction to bad earnings news. *Contemp. Account. Res.* 31, 911–936. doi: 10.1111/1911-3846.12050

Du, W. (2020). News and market efficiency in the Japanese stock market. *J. Behav. Finance* 22, 306–319. doi: 10.1080/15427560.2020.1774886

Eachempati, P., and Srivastava, P. R. (2021). Accounting for unadjusted news sentiment for asset pricing. *Qual. Res. Financ. Mark.* 13, 383–422. doi: 10.1108/QRFM-11-2019-0130

Fischer, T. (2011). News reaction in financial markets within a behavioral finance model with heterogeneous agents. *Algorithmic Finance* 1, 123–139. doi: 10.3233/af-2011-010

Grewal, J., Riedl, E. J., and Serafeim, G. (2019). Market reaction to mandatory nonfinancial disclosure. *Manag. Sci.* 65, 3061–3084. doi: 10.1287/mnsc.2018.3099

Hakkio, C. S., and Pearce, D. K. (1985). The reaction of exchange rates to economic news. *Econ. Inquiry* 23, 621–636. doi: 10.1111/j.1465-7295.1985.tb01786.x

Huang, B., Chen, L., and He, L. (2020a). How can government support affect behaviors of investors and rating agencies in a corporate bond market? Evidence from China's corporate bond market. *Emerg. Markets Finance Trade* 56, 485–507. doi: 10.1080/1540496X.2019.1651286

Huang, Z.-j, Huang, H., Song, Y.-y, and Feng, T.-y (2020b). Earnings management, analyst forecasts and credit rating of corporate bond: empirical evidences from Chinese listed companies. *J. Finance Econ.* 8, 21–32. doi: 10.12691/jfe-8-1-4

Kahneman, D., and Tversky, A. (2013). "Prospect theory: an analysis of decision under risk," in *Handbook of the Fundamentals of Financial Decision Making*, Part I, eds L. C. MacLean and W. T. Ziemba (Singapore: World Scientific Publishing Co. Pte. Ltd.), 99–127.

Kothari, S. P., Ramanna, K., and Skinner, D. J. (2010). Implications for GAAP from an analysis of positive research in accounting. *J. Account. Econ.* 50, 246–286. doi: 10.1016/j.jacceco.2010.09.003

Lin, K. J., Lu, X., Zhang, J., and Zheng, Y. (2020). State-owned enterprises in China: a review of 40 years of research and practice. *China J. Account. Res.* 13, 31–55. doi: 10.1016/j.cjar.2019.12.001

Loughran, T., McDonald, B., and Pragidis, I. (2019). Assimilation of oil news into prices. *Int. Rev. Financ. Anal.* 63, 105–118. doi: 10.1016/j.irfa.2019.03.008

Lu, H.-M., Tsai, F.-T., Chen, H., Hung, M.-W., and Li, S.-H. (2012). Credit rating change modeling using news and financial ratios. *ACM Trans. Manag. Information Syst.* 3, 1–30. doi: 10.1145/2361256.2361259

Matsubara, Y., Sakurai, Y., Prakash, B. A., Li, L., and Faloutsos, C. (2012). "Rise and fall patterns of information diffusion: model and implications," in *Proceeding of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, (New York, NY: ACM), doi: 10.1145/2339530.2339537

Narayan, P. K., and Bannigidadmath, D. (2017). Does financial news predict stock returns? New evidence from Islamic and non-Islamic stocks. *Pacific Basin Finance J.* 42, 24–45. doi: 10.1016/j.pacfin.2015.12.009

Partnoy, F. (2006). *How and Why Credit Rating Agencies Are Not Like Other Gatekeepers*. Available online at: https://ssrn.com/abstract=900257 (accessed February 15, 2022).

Qian, N. (2018). Anti-corruption effects on the credit risk of local financing vehicles and the pricing of Chengtou bonds: evidence from a quasi-natural experiment in China. *Finance Res. Lett.* 26, 162–168. doi: 10.1016/j.frl.2018.01.001

Ralston, D. A., Terpstra-Tong, J., Terpstra, R. H., Wang, X., and Egri, C. (2006). Today's state-owned enterprises of China: are they dying dinosaurs or dynamic dynamos? *Strateg. Manag. J.* 27, 825–843. doi: 10.1002/smj.545

Sabherwal, S., Sarkar, S. K., and Zhang, Y. (2011). Do internet stock message boards influence trading? evidence from heavily discussed stocks with no fundamental news. *J. Bus. Finance Account.* 38, 1209–1237. doi: 10.1111/j.1468-5957.2011.02258.x

Schwert, G. W. (1981). The adjustment of stock prices to information about inflation. *J. Finance* 36, 15–29. doi: 10.1111/j.1540-6261.1981.tb03531.x

Shepperd, J. A., and McNulty, J. K. (2002). The affective consequences of expected and unexpected outcomes. *Psychol. Sci.* 13, 85–88. doi: 10.1111/1467-9280.00416

Stracca, L. (2004). Behavioral finance and asset prices: where do we stand? *J. Econ. Psychol.* 25, 373–405. doi: 10.1016/S0167-4870(03)00055-2

Suleman, M. T. (2012). Stock market reaction to good and bad political news. *Asian J. Finance Account.* 4, 299–312. doi: 10.5296/ajfa.v4i1.1705

Tang, Y., Miao, X., Zang, H., and Gao, Y. (2018). Information disclosure on hazards from industrial water pollution incidents: latent resistance and countermeasures in China. *Sustainability* 10:1475. doi: 10.3390/su10051475

Thaler, R. H. (2019). The end of behavioral finance. *Financ. Anal. J.* 55, 12–17. doi: 10.2469/faj.v55.n6.2310

Tsai, F.-T., Lu, H.-M., and Hung, M.-W. (2010). "The effects of news sentiment and coverage on credit rating analysis," in *Proceedings of the Pacific Asia Conference on Information Systems*, (Taipei: PACIS).

Tsai, F.-T., Lu, H.-M., and Hung, M.-W. (2016). The impact of news articles and corporate disclosure on credit risk valuation. *J. Banking Finance* 68, 100–116. doi: 10.1016/j.jbankfin.2016.03.018

Xia, H. (2014). Can investor-paid credit rating agencies improve the information quality of issuer-paid rating agencies? *J. Financ. Econ.* 111, 450–468. doi: 10.1016/j.jfineco.2013.10.015

Zhao, H. (2006). Expectations of recruiters and applicants in large cities of China. *J. Manag. Psychol.* 21, 459–475. doi: 10.1108/02683940610673979

Zhao, L.-T., and Zeng, G.-R. (2019). Analysis of timeliness of oil price news information based on SVM. *Energy Proc.* 158, 4123–4128. doi: 10.1016/j.egypro.2019.01.821

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Zhang, Wang and Tong. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers | Frontiers in Communication

ORIGINAL RESEARCH
published: 11 April 2022
doi: 10.3389/fcomm.2022.855177

Pleasantness and Wellbeing in Poem Declamation in European and Brazilian Portuguese Depends Mostly on Pausing and Voice Quality

*Plinio A. Barbosa**

Department of Linguistics, University of Campinas, Campinas, Brazil

This work investigates the relationship between the sensations of pleasantness and wellbeing and the acoustics of poem declamation in two varieties of Portuguese: European and Brazilian. Ten speakers in each variety recited Alberto Caeiro's poem "Quando vier a primavera" to which dataset the declamation by two actors was added. Ten listeners in each variety participated in three perception tests for evaluating the degree of pleasantness and wellbeing of the entire reading in two different sessions by using a 5-degree Likert scale, and for choosing the more pleasant recitation from two recitations of a chunk of the poem. Both reciters and listeners were balanced in gender. A set of 22 prosodic parameters was used for predicting the listeners' evaluations in logistic regression models and linear discriminant analysis. Results showed that the parameters related to pausing and voice quality are the main explanatory variables with a minor role for the F0 median. The actors' recitations are preferred and gender differences are related to reciters, not listeners. Minor differences across the varieties were found.

Keywords: poem recitation, prosody, pleasantness, wellbeing, acoustic phonetics

OPEN ACCESS

Edited by:
Oliver Niebuhr, University of Southern Denmark, Denmark

Reviewed by:
Albert Rilliard, Université Paris-Saclay, France
Pärtel Lippus, University of Tartu, Estonia

***Correspondence:**
Plinio A. Barbosa
pbarbosa@unicamp.br

Specialty section:
This article was submitted to Language Sciences, a section of the journal Frontiers in Communication

Received: *14 January 2022*
Accepted: *11 March 2022*
Published: *11 April 2022*

Citation:
Barbosa PA (2022) Pleasantness and Wellbeing in Poem Declamation in European and Brazilian Portuguese Depends Mostly on Pausing and Voice Quality. Front. Commun. 7:855177. doi: 10.3389/fcomm.2022.855177

INTRODUCTION

The vocal differences between speaking styles are related to changes in voice quality, as well as rhythm and intonation, according to studies carried out in the field for decades (see Eskénazi, 1993 for a review). Among these, professional styles such as TV and radio broadcasting (Léon, 1993), and reading in literary (book) festivals common in all Europe are remarkable for provoking effects related to pleasantness and wellbeing in the listeners (Fónagy, 1961, 2001). As regards poetry reading, although the aesthetic effect is proper to the culture in which the poetic texts were born, it also depends on prosody, with elements such as novelty and complexity being associated with prosodic aspects such as the rhythmic and intonation organization of the verses that have effects of surprise and expectation, among others (Jacobsen, 2010). A study by Tegnér et al. (2009) showed that other effects can include the increase of wellbeing in cancer patients after 6 weeks of poetry reading groups by causing an improvement of their emotional resilience and a decrease of their anxiety levels.

In a review of studies on the topic of attractiveness in speech, Rosenberg and Hirschberg (2020) showed that, although the concept of "attractiveness" is very context-dependent, common points can be found, such as the fact that male voices with a low fundamental frequency (F0) are preferred by women and are associated with dominance, whereas female voices with a

higher F0 are the preference of men. Quené (2020), Strangert and Gustafson (2008) and Hodges-Simeon et al. (2010) have reported that women also prefer men who have higher speech rates. Baumann (2017) found that women generally found the speech of other women more pleasant, but this result is controversial considering the results by Burkhardt et al. (2010), which suggested this is not true.

Work on German showed that attractiveness, regardless of gender, seems to increase with the variability of F0 (Weiss et al., 2020) as well as with the absence of disfluencies (Strangert and Gustafson, 2008; Weiss and Burkhardt, 2012). Weiss and Burkhardt (2010) also found a relationship between soft and breathy speech, and a lower spectral center of gravity with the growth of positive appreciation by listeners, a result in the same line as the one by Carlsen et al. (2018) who showed that lower harmonic-to-noise ratio (HNR) and higher jitter are related to an increase in pleasantness. As regards creakiness, Anderson et al. (2014) found that the voices of young adult females sound less attractive, less competent, and less educated when they read a phrase imitating vocal fry (creaky voice).

As regards work on Portuguese, pleasantness and aesthetical appreciation of poem recitation were lines of investigation in both European (EP) and Brazilian Portuguese (BP). The main interest of the works by Braga et al. (2007) and Pinto-Coelho et al. (2013) was the choice of human voices for EP Text-to-Speech (TTS) systems. In the first work, the authors evaluated eight subjective dimensions, including pleasantness, in a test with 2-min audio excerpts of 62 female voices reading texts of information delivery aiming at sounding talented as a speaker (these women certainly intended to sound pleasant given that instruction). They showed that pleasantness is one of the components of the choice for a particular voice, but that the general preference was for (female) voices with low fundamental frequencies. In the second work, the authors built a machine learning system for classifying and predicting the degree of the pleasantness of 3-min audio excerpts from 77 EP-speaking female speakers. The training was designed to predict a previously web-based evaluation of a 5-point pleasantness scale obtained from 112 participants from prosodic- and voice-quality acoustic parameters. Results showed that jitter, shimmer, F0 mean, derivative of intensity, and maximum of F0 derivative were the best predictors of pleasantness.

As regards the effects of poetry recitation, by working on BP, Madureira et al. evaluated the declamation of a single poem at each time by one or more reciters, both professional and non-professional. Madureira (2008) investigated the perceptual effects of the production choices of two professional speakers (male and female) reciting the same poem, "Soneto da Fidelidade" from Vinícius de Morais. A group of 30 listeners evaluated the impressions carried out by the voices in both recitations. That of the male speaker was considered enthusiastic and splendorous, whereas that of the female speaker evoked an impression of grief and sadness. The acoustic analysis of both readings revealed that whispery voice, F0 narrow range, and a great number of silent pauses were the production strategies used by the female speaker to express those negative effects, whereas the male speaker employed much varied intonation patterns to sound more expressive to provoke a sensation of liveliness. Madureira and Fontes (2019) analyzed face movements and acoustic parameters of the recitation of the poem "A Valsa" by a Brazilian actor to show that LTAS slope, mandibular range, and F0 maximum are related to surprise and contempt inferred automatically from the Affectiva facial analysis system. It was also a single professional speaker reading a single poem that was analyzed by Madureira and Camargo (2010) to show a relation between prosodic parameters and subjective dimensions. These dimensions were evaluated by 30 college students in a semantic differential scale questionnaire with the following descriptors: activation (calm/activated), valence (pleasant/ unpleasant), emotion (joy, sadness, anger, exasperation), and speech acts (advice, admonition, order, and plea). Repetitions of verses recited in different ways were used to relate prosodic variation and perceptual effect. They showed that pleasantness was related to an expanded pharynx, which causes F0 lowering (see Camargo et al., 2019). In a follow-up study of the effects of the recitation of the poem "Soneto da Fidelidade" (Menegon et al., 2021), this time recited by eight non-professional speakers, pleasantness was moderately relevant as a subjective parameter and related to voice quality and the parameter of LTAS slope. These works on BP and EP suggest that melodic, temporal, and voice quality dimensions could explain subjective positive evaluations in both poetry (BP) and ordinary prose (EP).

As can be seen from this overview, there are issues related to the link between prosodic-acoustic parameters and the sensations of attractiveness, pleasantness, and speech positive appraisal that seem to be recurrent, while others need further study such as the question of appreciation by men and women of the same-gender vs. cross-gender speech. Further investigation on acoustic parameters such as correlates of vocal effort, pausing, intensity variability could shed light on aesthetic appreciation. These measures are included here.

Besides extending the spectrum of parameters related to the acoustics of pleasant speech, additional research on Portuguese could help contribute to advance knowledge on the link of speech production and subjective appreciations such as pleasantness and wellbeing, the latter an under-studied theme.

Poetry seems to be appropriate for examining this link because declamation usually focuses on provoking a particular effect on the listener, and not only on delivering information. It is worth noting that the scientific community is developing a sense of the importance of poetry as a means of increasing quality of life and this can be exemplified with a journal specializing on the results of therapy using poem reading and declamation, the *Journal of Poetry Therapy* sponsored by the National Association for Poetry Therapy in the USA and publishing since 2003. In this vein, it can be said that poetry actually seems to increase wellbeing in a therapy situation, at least. That is why the present work investigates the link between the acoustics and the perception of pleasantness and wellbeing of poem recitation in two varieties of Portuguese: Brazilian and European. The findings of this work can also shed light on the relation between the use of voice and speech and aesthetical pleasure, which give elements for the training of professional speakers.

The work presented here concerns the evaluation of the degrees of the listener's feelings of pleasantness ("agradabilidade" in Portuguese) and wellbeing ("bem-estar" in Portuguese) caused by the listening of a Portuguese poem, as explained in detail in Section Pleasantness and Wellbeing Evaluations. Similarly, to other works reviewed here, this investigation also uses the recitation of a single poem but expands the number of speakers to twenty, which allows a larger study of prosodic variability.

Based on the examined literature and general expectations based on cultural aspects or common sense, the main hypotheses for poem recitation in this study are:

- H1. Pleasantness and wellbeing appraisals are equivalent, but the second has a weaker relation to acoustics because it is more subjective than the first;
- H2. Lower mean F0 increases both the sensations of pleasantness and wellbeing for women listening to the recitation of men;
- H3. Higher mean F0 increases both the sensations of pleasantness and wellbeing for men listening to the recitation of women;
- H4. Diverging from what is said in the literature on attractiveness, due to the appraisal of care with the content in poem recitation, slowing down and more frequent pauses increase both pleasantness and wellbeing;
- H5. Breathier speech increases pleasantness and wellbeing;
- H6. More variable F0 increases pleasantness and wellbeing;
- H7. The lesser the vocal effort, the higher the sensations of pleasantness and wellbeing;
- H8. Portuguese listeners appreciate the declamation by Brazilians more than Brazilians appreciate the declamation by Portuguese speakers, due to the greater exposure in Portugal to the speech of Brazilians in Portuguese TV, especially Brazilian soap operas;
- H9. The acoustic parameters related to pleasantness and wellbeing are the same and in the same degree of relative importance in the two varieties of Portuguese, in the absence of evidence to the contrary;
- H10. Men and women use the same prosodic-acoustic parameters to make the speech better appreciated by the listener;
- H11. Actors' voices are more pleasant and cause more wellbeing than those of lay reciters.

In the next section, the methodology for testing the eleven hypotheses above is presented. Section Results sets forth the results, which are discussed in detail in Section Discussion. Section Conclusion summarizes the main findings.

METHODOLOGY

The PROS-POIESIS corpus is formed by the declamation of the poem "Quando vier a primavera" (When Spring comes) by Alberto Caeiro, one of Fernando Pessoa's heteronyms, by ten Brazilian speakers and ten Portuguese speakers in balanced gender. All speakers do not have voice professional training and have between 25 and 50 years of age to avoid an age group with effects of vocal aging (Stathopoulos et al., 2011) that usually affect the values of parameters such as F0 median, jitter, shimmer, and speech rate.

Due to the Covid-19 pandemic, the participants themselves used the Easy Voice app on their cell phones to make all recordings. Because this app allows choosing among different codifications, instructions were given to record all audio files in PCM format (WAV) at a sampling rate of 48 kHz. The author, who is a trained phonetician, further evaluated all audio files. The recordings were resampled at 16 kHz and leveled to the same maximum intensity level at 65 dB.

As control recordings, the declamations of the same poem by Brazilian actor Ivan Lima (72 years old at the time of recording) and by Portuguese actor Pedro Lamares (33 years old at the time of recording) were included in the dataset. The reason for choosing these two actors was simply the public availability of these recordings for the same poem. It is expected that they get the highest scores of pleasantness and wellbeing due to their professional voices. The recitation by Pedro Lamares is available here since 2011: <https://www.youtube.com/watch?v=ZNWEmKLLFWA> and that by Ivan Lima, here since 2020: <https://www.youtube.com/watch?v=7QP_cEEQ8RA>. Unfortunately, no actresses were found reading the same poem in both varieties of Portuguese.

Acoustic Parameters

The Prosody Descriptor Extractor script for Praat (Boersma and Weenink, 2021) implemented by the author (Barbosa, 2020) was used to extract 22 prosodic-acoustic parameters from 14 chunks from one to two verses of the spoken poem including the silent pause at the end, when applicable. This segmentation is shown in the Appendix in the original version followed by a translation for the sake of making the meaning transparent for the general reader. The criteria used for choosing this segmentation are two-fold: (1) to ensure an acceptable unit of meaning when each chunk is heard in isolation; (2) to ensure a minimum duration of about 3 s for the computation of the acoustic parameters in each chunk as explained in the following paragraphs.

The segmentation of the chunks was the same across the 20 reciters and the reading of the poem lasted from 60 to 95 s. A set of 22 parameters was computed for each chunk, a single value per chunk. This set can be divided into the following classes: (1) Class melody: twelve descriptors of fundamental frequency (F0): median, semi-amplitude between quartiles (F0SAQ), F0 minimum and maximum, standard-deviations of values and time positions of F0 local peaks, F0 peak rate, and F0 mean peak bandwidth within each chunk, mean and standard-deviations of rates of F0 rises and falls; (2) Class intensity: two intensity descriptors: spectral emphasis and the coefficient of variation of global intensity (the ratio between global intensity standard-deviation and mean); (3) Class VQ: four voice quality descriptors: long-term averaged spectrum (LTAS) slope, Harmonic-to-Noise Ratio (HNR), jitter and shimmer; finally, (4) Class timing: four temporal descriptors: silent pause average duration, pausing rate, speech, and articulation rates.

Melodic measures such as the statistical descriptors of centrality and dispersion (median and semi-amplitude between

quartiles) and measures of extreme values (F0 minimum and maximum) are more easily found in the literature of prosodic variation across styles. To them, we added parameters aiming at describing other aspects of melodic variability. Those are the descriptors related to the peaks of F0 within a chunk. These peaks are automatically obtained by the Prosody Descriptor Extractor script after a three-step technique for obtaining a smoothed F0 contour in Praat: first, a 5-Hz low-pass filter applied to the extracted F0 contour; second, a quadratic interpolation and then a Praat-built function to select the extreme values. From these values, we computed four peak-related measures: the standard-deviation of the F0 peak values, the standard deviation of the F0 peak time positions, the mean of F0 peak rate, and the mean of F0 peak bandwidth. The first three of these F0 peak-related measures are connected to liveliness because it concerns how F0 peaks are distributed across a sound excerpt. The majority of these F0 peaks are pitch accents because of the care to use a 5-Hz smoothing function to highlight the main F0 peaks. Since the mean of F0 peak rate is about 1 s in our data, a minimum duration of 3 s for a chunk allows the computation of mean and standard deviations of F0 peak intervals, which is useful for further investigating F0 peak rate variability across the chunks.

As for F0 bandwidths, by studying the possible correlates of charismatic speech in the speeches of Steve Jobs (Apple) and Mark Zuckerberg (Facebook, recently renamed Meta), Niebuhr et al. (2018) have shown, among other aspects, that Jobs has a broader mean pitch accent bandwidth than Zuckerberg, which possibly contributes to the former sound more charismatically. Because charisma is a positive evaluation that can be attributed to listening to an audio file, we decided to include mean F0 peak bandwidth as a possible parameter for explaining the positive dimensions of pleasantness and wellbeing.

Spectral emphasis was computed according to Traunmüller and Eriksson (2000) as the difference between the spectral energy of the whole band (E) and that of the low band with a cut-off frequency of 400 Hz (E_0) for computing the energy of F0. This cut-off frequency is an approximation of the original proposal of 1.5 times the mean fundamental frequency, which has the same goal of computing the energy of F0 in the lower band. According to the authors, the higher the value of spectral emphasis, the higher the vocal effort. In previous works, we computed spectral emphasis in intervals as short as vowel intervals and showed that this measure was one of the three relevant ones (with duration and F0 standard deviations) to signal levels of lexical stress in Brazilian Portuguese and Swedish (Barbosa et al., 2013; Eriksson et al., 2013). That is why its computation within a 3-s minimum duration chunk is robust enough for the analysis presented here.

LTAS slope, computed here by the difference in mean energy between the bands 0–1 and 1–4 kHz, varies according to differences in voice quality for a series of reasons including breathiness (the flatter the slope, the breathier the voice, see Mendoza et al., 1996), differences in voice projection for singing and acting (Sundberg, 1994), and other aspects not explored here that can be found in the review by Master et al. (2006). This parameter was included to evaluate its potential in explaining the listeners' preferences. A word of caution is necessary though because authors such as Löfqvist and Mandersson (1987) have shown that LTAS measures stabilize after 10 s of continuously voiced frames of speech, due to variability in articulation. Nevertheless, we found in our data that its variability across the chunks of the same reciter was lesser than its variability across reciters, which can be used as a criterion of discriminability for this parameter across the recitations and a potential predictor of the listeners' preferences for a particular recitation.

Harmonic-to-Noise Ratio (HNR) is a correlate of roughness and breathiness (de Krom, 1994; Verdonck-de Leeuw et al., 2001) as well as creakiness (Khan et al., 2015): the lower its value, the rougher, the breathier and the creakier the speech is. Although often computed from sustained vowels, it has been used in connected speech with success for predicting roughness (Severin et al., 2005). Because low values of HNR can be associated with distinct qualities such as creaky and breathy voices, this measure should be combined with others, such as LTAS, jitter, and shimmer, to discriminate between the specific voice quality.

Jitter and shimmer are measures of irregularity across glottal cycles, the former of glottal period irregularity and the latter of glottal amplitude irregularity. Both are usually computed for sustained vowels but their computation in connected speech could reveal differences in voice qualities because, according to Laver (1980), voice quality is "the characteristic auditory coloring of an individual speaker's voice [...] a cumulative abstraction over a period of time of a speaker-characterizing quality."

Pleasantness and Wellbeing Evaluations

Five perception tests were carried out, where the first four consisted in the attribution of the degree of pleasantness or wellbeing felt by the listeners, according to the instruction given to them: (1) one Likert-scale test for evaluating pleasantness applied to 10 Brazilian listeners and containing recitations by Brazilian speakers; (2) one Likert-scale test for evaluating wellbeing applied to the same 10 Brazilian listeners and containing recitations by Brazilian speakers; (3) one Likert-scale test for evaluating pleasantness applied to 10 Portuguese listeners and containing recitations by Portuguese speakers; (4) one Likert-scale test for evaluating wellbeing applied to the same 10 Portuguese listeners and containing recitations by Portuguese speakers; and (5) a discrimination test applied to all Brazilian and Portuguese listeners for evaluating paired chunks from both European (EP) and Brazilian Portuguese (BP) recitations. All instructions were given in Portuguese.

As part of the four Likert-scale tests carried out with Brazilian and Portuguese listeners, we included the declamation of the actors to the ones of the ten lay speakers in the data set of each variety of Portuguese. This allowed comparing the evaluations of the actors with those of the lay reciters to answer Hypothesis 11.

The scale for pleasantness evaluation varied in five degrees from "very unpleasant" ("muito desagradável" in Portuguese, degree 1) to "very pleasant" ("muito agradável" in Portuguese, degree 5) with the neutral response ("neutro" in Portuguese) having degree 3. Similarly, the scale for wellbeing evaluation varied in five degrees from "very low wellbeing" ("grande mal-estar" in Portuguese, degree 1) to "very high wellbeing" ("grande bem-estar" in Portuguese, degree 5) with the neutral response also having degree 3. Because the labels have general use in

Portuguese for evaluating external events as being pleasant ("agradável") or causing wellbeing ("bem-estar"), no particular instructions were given to the participants on how to interpret them. Although the term "mal-estar" could also be interpreted as related to illness, because it was presented in a scaling factor containing the term "bem-estar," the possibility of interpreting "mal-estar" as causing a sensation of illness to the listener must be discarded. The similarity of the responses for both pleasantness and wellbeing strongly suggests that this illness-related interpretation was not considered by the listeners in the context of the effect of a poem recitation.

As for the discrimination test, a set of 44 pairs using the 14 chunks of the poem recitations separated by a short musical tone and combining different genders and speakers was prepared. The chunks being compared in each pair have the same content and are from the same language variety for avoiding a choice biased toward differences in content or variety. The pool of paired audio files contains all chunks of the poem presented from 2 to 6 times from different reciters to combine readings from the same gender (either males or females) and across genders, taking at least a chunk from all reciters. Apart from that, the choice of a particular pair of reciters in an audio file pair was random. To evaluate the behavior of listeners according to variety, the pooled group of 20 listeners from Brazil and Portugal evaluated the same set. Their task was to say which of the readings is the one that sounds more pleasant to him/her, the first or the second.

For the five tests, the listeners were selected in the same age range as the reciters, that is, between 25 and 50 years of age, to avoid the effects of differences in evaluation from younger or older age groups. The discrimination test also allowed checking the possible effect of the variety of the reciter and the listener when evaluating the pairs of chunks in both BP and EP, as well as whether the mean degrees attributed in the Likert scale tests could predict the performance of the listeners in the discrimination test. Each listener had online participation in the three tests, which were built in the Survey Monkey platform with the help of a research assistant. The discrimination test was performed after the Likert scale tests and the three tests were completed in 30–45 min.

Statistical Models

To test our hypotheses, three kinds of statistical models were run in the R (R Development Core Team, 2008) software. The first kind were correlation tests applied to verify if Brazilian and Portuguese listeners answered in a similar way to the stimuli of the discrimination test and if these responses are correlated with the differences in degree from the Likert-scale test. The correlation for evaluating the degree of similarity of the responses for pleasantness and wellbeing by the listeners of the two varieties was computed to answer Hypothesis 1.

The second kind of statistical model was logistic regression. After a process of progressive model simplification, four final logistic regression models, corresponding, respectively, to the four Likert-scale tests presented in the previous section, were retained that predicted the degree of pleasantness or wellbeing (one of the two predicted variables in each model) from the corresponding prosodic-acoustic parameters (the predictor variables). For doing so, the degrees from 1 to 5 were, respectively, transformed to 0 to 100% in intervals of 25%, allowing the use of a logistic model. The quasi-binomial family in the R *glm* function was used for prediction when there were important differences between degrees of freedom and amount of deviance, according to statistical theory (Roback and Legler, 2021, chapter. 6). When this was not the case, the binomial family was used. Pseudo-correlation measures were used to evaluate the degree of explained variance of these models, which are measures of effect size. Model simplification always started by using non-correlated acoustic parameters as predictors and successively eliminating those that were not significant (see below for the process of eliminating correlated parameters). The final models contain only significant predictors.

The third kind of statistical model was a model based on linear discriminant analysis (LDA). The final model predicts one of the two possible responses of the corresponding discrimination test (first or second chunk) from the differences of the mean values of each parameter from the two recitations of the respective chunk. In the LDA analysis, there were no predictors related to pausing (pause duration and rate of pause) because a single silent pause ends the majority of the chunks, and this final pause coincides with the end of an utterance, then, pause duration cannot be apprehended. The process of simplification also considered the maximum number of non-correlated predictors with a progressive elimination of non-significant predictors for retaining only significant predictors in the final model.

For all models, when two or more predictors of the same class correlated by more than 50%, one of them was taken as a predictor and the other one(s) was not used in the initial model. That procedure allowed us to discard the following parameters: F0 minimum and maximum (both correlated to more than 60% with F0 median), standard-deviations of rates of F0 rises and falls (both correlated to more than 60% with F0SAQ), and speech rate (correlated with more than 70% with articulation rate). In all the statistical models built, the 5% level of significance was chosen.

All stimuli are publicly available for listening at the following URL: <https://figshare.com/authors/Plinio_Barbosa/11320902>.

RESULTS

Figure 1 presents the medians of the degrees attributed to the scales of pleasantness and wellbeing according to the variety and gender of the groups of listeners. No differences between the evaluations of male and female listeners were observed when checking the closeness of median responses either for male or female reciters. The correlation between the evaluation of the two scales was 82% for Brazilian listeners and 89% for Portuguese listeners. This confirms the first part of Hypothesis 1, the close relationship between the two scales as well as the fact that the left part ("mal-estar") of the wellbeing scale was not interpreted differently from the right part ("bem-estar"), otherwise, the correspondence between the two scales would be low.

In the logistic models, the gender of the reciter in each spoken variety happened to produced different sets of predictors,

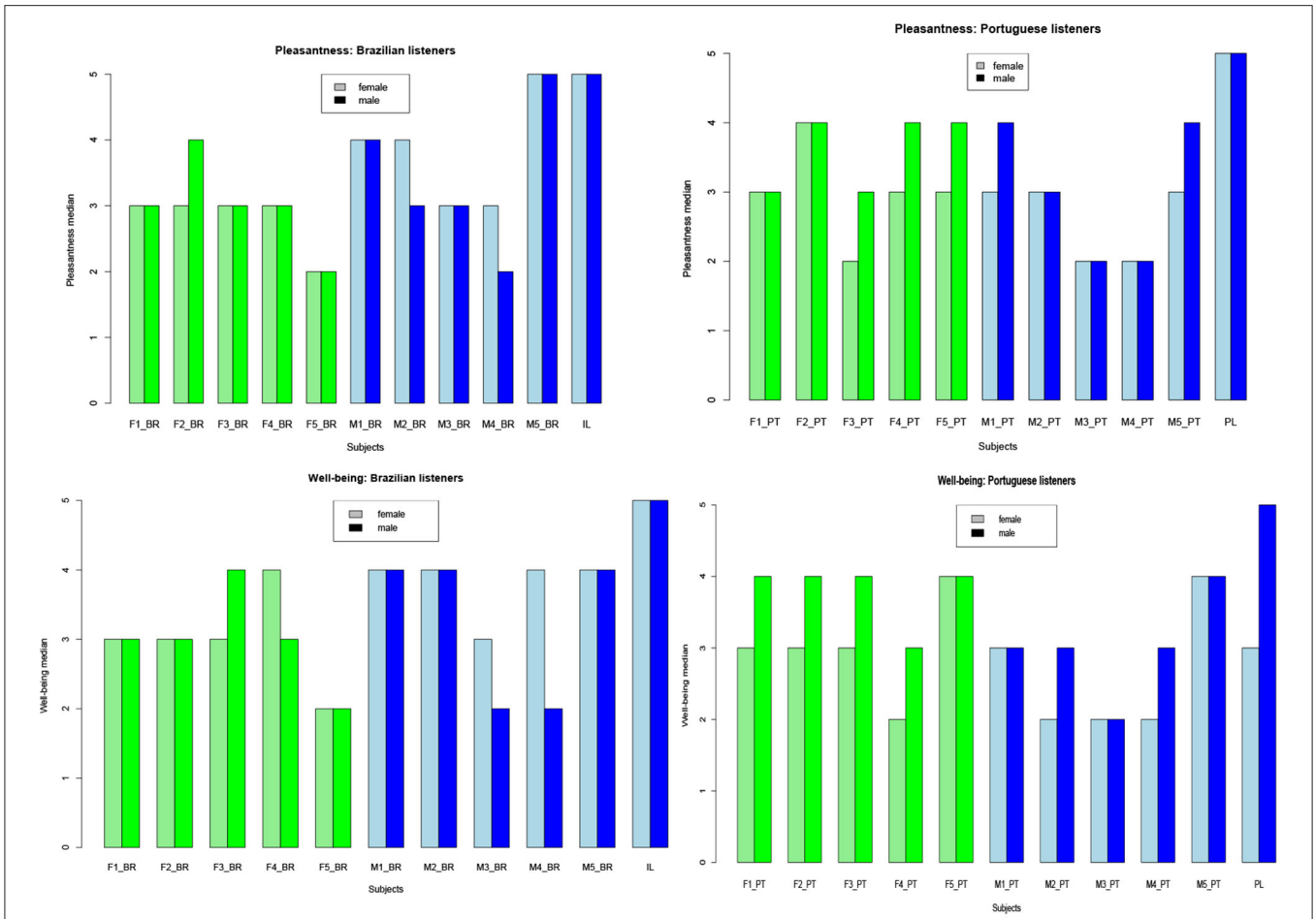


FIGURE 1 | Medians of degrees attributed when Brazilian (left) and Portuguese (right) listeners evaluated the pleasantness (top) and the wellbeing (bottom) of the whole poem. The gender of the reciter is indicated (Mn signals a male reciter, Fn signals a female reciter n). IL is the Brazilian actor and PL the Portuguese actor. The gender of the listener is indicated by the darkness of the color (the darker bar to the right in the contiguous bars are from male listeners).

that is why each one of the four original models was split in two according to the reciter's gender and these results are then presented gender-wise for each language variety. For the datasets of the two varieties of Portuguese, if a parameter in a specific chunk could not be computed for mathematical reasons, the measures of all the other parameters were not included in the models, for maintaining the whole set of predictors in each model. An example of this case is the computation of peak rate, which requires at least two F0 peaks in the chunk. That is why some boxes in **Figures 2**–**5** do not present any variability but a median value only. For pleasantness, the actors in each community got the 5 median degrees, whereas for wellbeing the Portuguese actor (PL) got the 4 median degrees and the Brazilian actor (IL) got the 5 median degrees (see **Figure 1**).

The degree of pleasantness attributed to Brazilian female reciters by the Brazilian listeners is explained by three parameters in order of importance: rate of pause (the lower, the more pleasant), the slope of LTAS (the faster the slope, the more pleasant), and pause duration (the longer, the more pleasant). Pseudo-correlation explains 42% of the variance. The degree of pleasantness attributed to Brazilian male reciters by the same listeners is explained by four parameters: the same as for female reciters in the same order of importance plus F0 median in the last position (the higher, the more pleasant). The parameters explained 23% of the variance. The distribution of the four parameters' values of the BP dataset that are the predictors of the two models can be checked graphically with the boxplots in **Figure 2**. In the case of the rate of pause (top, right) the tendency is not linear and higher means are found for degrees 3 to 5 (recall that the bold horizontal segment in boxplots stands for the median, not the mean).

As for the degree of pleasantness attributed to Portuguese female reciters by the Portuguese listeners, only pause duration was significant (the longer, the more pleasant), explaining 44% of the variance. As for the effect of male recitations on the same listeners, three parameters in order of importance explain the listeners' choices: spectral emphasis, F0 median (the lower, the more pleasant in both cases), and pause duration (the longer, the more pleasant), which explains 38% of the variance. The distribution of the parameters' values that are the predictors of

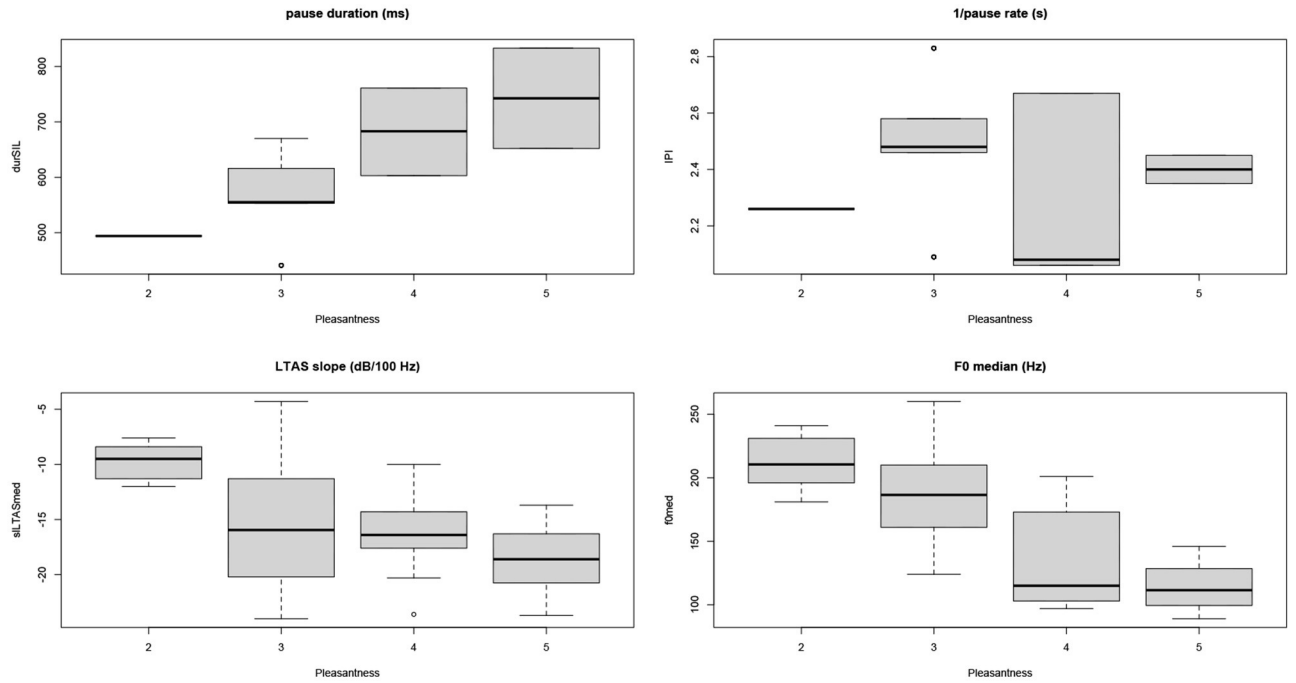


FIGURE 2 | Boxplots of the four acoustic parameters that predict pleasantness for the BP dataset.

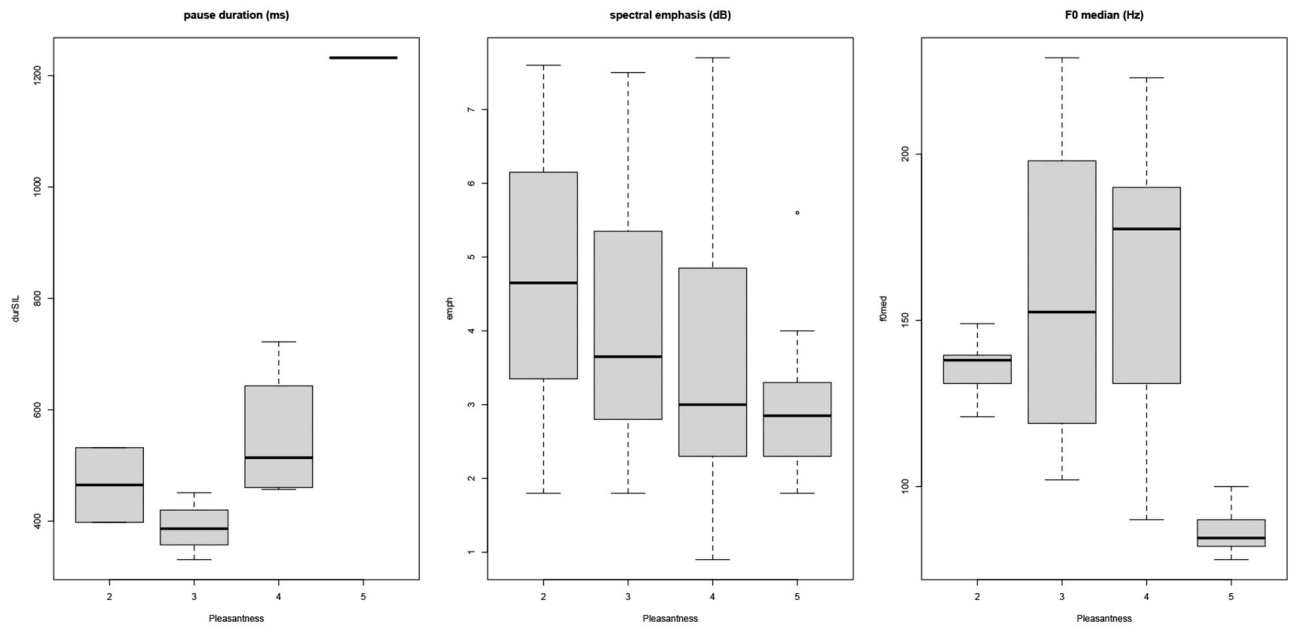


FIGURE 3 | Boxplots of the three acoustic parameters that predict pleasantness for the EP dataset.

the two models can be checked graphically with the boxplots for the three predictors for the EP dataset in **Figure 3**.

The degree of wellbeing attributed to Brazilian female reciters by the Brazilian listeners has a very low explained variance, only 5%. Only two parameters explain the responses, in order of importance: HNR (the higher, the greater the wellbeing) and slope of LTAS (the faster the slope, the greater the wellbeing). By contrast, the degree of wellbeing attributed to Brazilian male reciters by the same listeners has 27% of explained variance, higher than the one from the model that evaluated pleasantness,

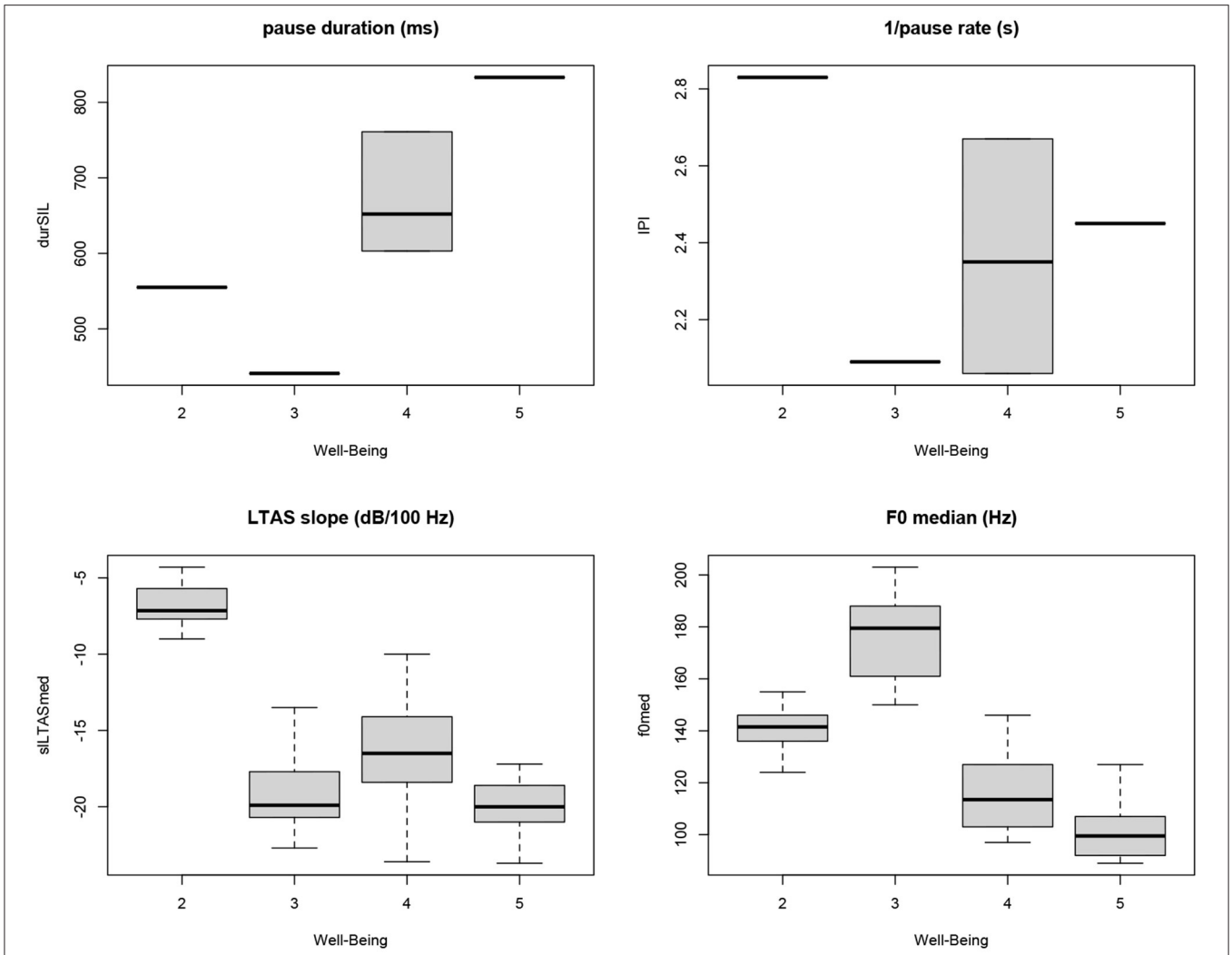


FIGURE 4 | Boxplots of the four acoustic parameters that predict for the BP dataset. Results for male reciters only.

and explained by the same four parameters: rate of pause (the lower, the higher the wellbeing), pause duration (the longer, the higher the wellbeing), the slope of LTAS (the faster the slope, the higher the wellbeing) and F0 median (the higher, the higher the wellbeing). **Figure 4** shows the boxplots for the parameters explaining wellbeing evaluation of the male recitations due to its higher explained variance. The same tendencies found for pleasantness prediction can be seen, although less gradual. Degree 3 is clearly the one that deviates from a linear trend for degrees 2, 4, and 5 in all four plots, and certainly contributed to the positive contribution of the F0 median to explain wellbeing (the higher, the higher the wellbeing).

As for the degree of wellbeing attributed to Portuguese female reciters by the Portuguese listeners, the same 5% of explained variance of the BP dataset evaluation by female listeners was found, but associated to different parameters: F0med (the higher, the higher the wellbeing) and pause duration (the longer, the higher the wellbeing). As for results for male reciters given by the same listeners, two parameters in order of importance explain the data: pause duration (the longer, the higher the wellbeing) and spectral emphasis (the lower, the higher the wellbeing), which explains 32% of the variance. **Figure 5** shows the boxplots for the parameters explaining wellbeing attributions for the male recitations only because it is the model with a higher explained variance where lower values for spectral emphasis and higher values for pause duration are associated with higher degrees of wellbeing.

The results presented above signal, as far as female reciters are concerned, that the relation of wellbeing to acoustic parameters is weaker than that for pleasantness. For male reciters, the relation of both scales to acoustic parameters is very similar, though. This partially confirms the second part of Hypothesis 1.

As for the discrimination results, the overall correlation between the responses of Brazilian and Portuguese listeners was 78%. When the variety of the reciter was BP, their correlation was 82 and 76% when the variety of the reciter was EP. When

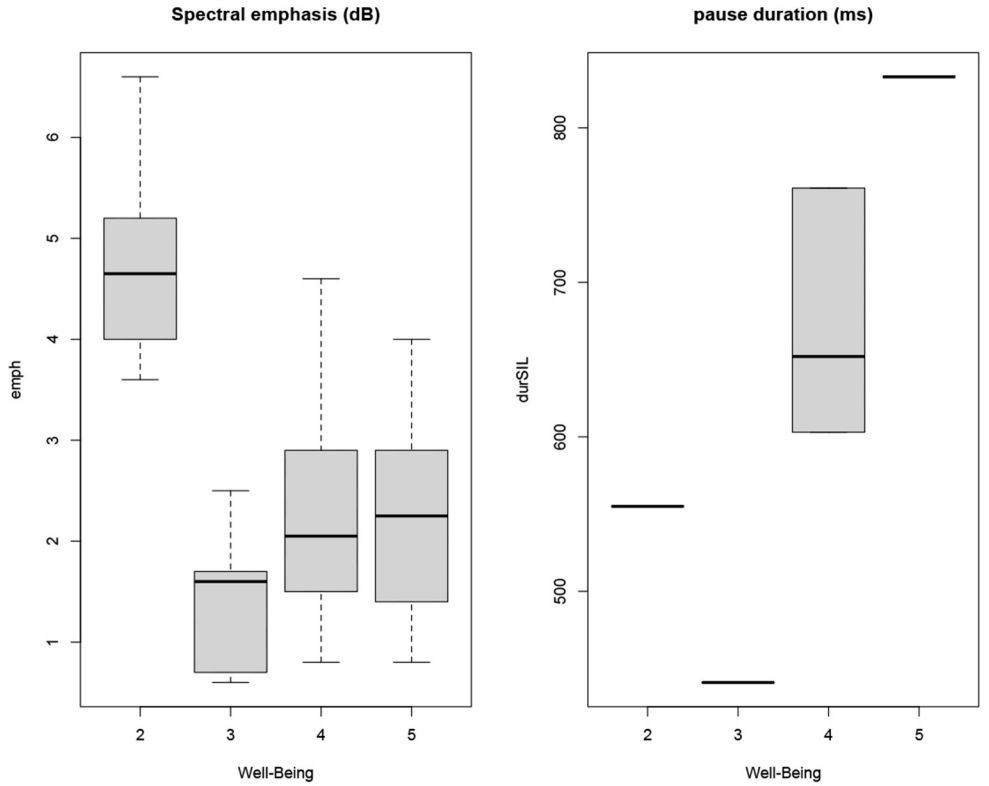


FIGURE 5 | Boxplots of the two acoustic parameters that predict wellbeing for the EP dataset. Results for male reciters only.

both recitations of a chunk were from male speakers, the two groups of listeners correlated 78%, in comparison with 83% when they evaluated female reciters only and 76% for pairs containing mixed-gender reciters. Prediction for first or second preferred chunk in terms of pleasantness got the same proportion of hits in the LDA model: 75%.

The overall prediction of the listeners' choices is explained by four parameters in order of importance in the final model: articulation rate (faster rates are preferred), F0 falls (smoother slopes are preferred), spectral emphasis (lower values, meaning softer readings are preferred), and having a lower effect, the slope of LTAS (sharper slopes are preferred, which is related to less effort). If the responses are separated according to the variety of the reciter, only the first three parameters of the overall model are significant in different orders of importance: the same of the overall model when Brazilian reciters were evaluated, with 88% of hits in both directions, and the following order when Portuguese reciters were evaluated: spectral emphasis (lower values are preferred), F0 falls (smoother slopes are preferred), and articulation rate (faster rates are preferred). The preference for the first reading was predicted with 78% of hits and 75% of hits for the second reading choice, signaling a balance of preference between the two positions.

The median response by both Portuguese and Brazilian listeners for the chunk read by the Brazilian actor was for his reading in 6 out of 6 comparisons with lay reciters, whereas the median response for the chunk read by the Portuguese actor was for his reading in 5 out of 6 comparisons with lay reciters. A Portuguese female, the lay reciter was preferred over the Portuguese actor in this single exception. This confirms an overall preference for the actors in the discrimination test as well. All statistical results are summarized in **Table 1**.

For the sake of examining the relationship between the acoustic parameters and the preferences for particular reciters, **Table 2** gives the mean values for the significant predictors considering into account the models involving Portuguese reciters. It can be seen that male reciters M5 and PL (the actor), who were more pleasant, have lower values for spectral emphasis and F0 median and higher pause duration means, which is in accordance with the logistic regression model. Female reciter F2 has the highest pause duration mean, and was also one of the females with high scores for pleasantness.

Although the LDA model evaluates a particular chunk, it can be seen that higher articulation rates and lower spectral emphasis are associated with higher scores of pleasantness.

Table 3, on the other hand, gives the mean values for the significant predictors considering into account the models involving Brazilian reciters. It can be seen that for male reciters M5 and IL (the actor), the ones who were more pleasant, have low rates of pause, the longest pauses among males, and fast LTAS slopes, which is in accordance with the prediction of the corresponding logistic regression model. Female reciter F2 has

TABLE 1 | Statistical models' summary. "Log. reg." stands for "logistic regression," and "pleasant." stands for "pleasantness."

Model type	Dataset	Language	Reciters' gender	Var. predicted	Model fitting	Predictor(s)
Log. reg.	Poem	BP	Female	Pleasant.	42%	Pause rate> LTAS slope > pause duration
Log. reg.	Poem	BP	Male	Pleasant.	23%	Pause rate> LTAS slope > pause duration > F0median
Log. reg.	Poem	BP	Female	Wellbeing	5%	HNR > LTAS slope
Log. reg.	Poem	BP	Male	Wellbeing	27%	rate of pause > pause duration > LTAS slope > F0 median
Log. reg.	Poem	EP	Female	Pleasant.	44%	Pause duration
Log. reg.	Poem	EP	Male	Pleasant.	38%	Spectral emphasis > F0 median > pause duration
Log. reg.	Poem	EP	Female	Wellbeing	5%	F0 median > pause duration
Log. reg.	Poem	EP	Male	Wellbeing	32%	Pause duration > spectral emphasis
LDA	Paired chunks	EP & BP	Male & female	Pleasant.	88%	Articulation rate > F0 fall > spectral emphasis > LTAS slope

Refer to Section Acoustic parameters for an explanation of acoustic parameters.

TABLE 2 | Mean values for significant predictors for Portuguese reciters (gender indicated by the first letter: F, female; M, male).

	F0 median (Hz)	F0 fall (Hz/frame)	Pause duration (ms)	1/pause rate (s)	Articulation rate (syll/s)	HNR (dB)	SE (dB)	LTAS slope (dB/100 Hz)	Pleasant
F1	210	−4.6	384	1.7	6.7	13.0	5.5	−13.5	3
F2	185	−3.2	564	2.2	6.1	9.3	5.8	−9.3	4
F3	185	−3.8	389	2.5	6.3	11.9	4.9	−13.1	3
F4	176	−4.2	457	1.9	7.3	9.5	3.8	−14.0	4
F5	191	−3.5	464	2.2	6.3	14.0	2.4	−17.4	4
M1	118	−3.1	331	1.6	7.5	7.4	2.9	−14.2	3
M2	120	−1.8	451	1.7	7.4	7.4	2.8	−16.2	3
M3	133	−1.7	532	2.3	6.0	7.6	6.2	−9.5	2
M4	139	−1.8	398	1.8	6.5	10.4	3.2	−11.8	2
M5	100	−2.2	722	2.2	6.7	5.7	2.3	−16.0	4
PL	86	−0.8	1,232	2.9	8.1	1.8	3.0	−11.9	5

PL is the Portuguese actor.

TABLE 3 | Mean values for significant predictors for Brazilian reciters (gender indicated by the first letter: F, female; M, male).

	F0 median (Hz)	F0 fall (Hz/frame)	Pause duration (ms)	1/pause rate (s)	Articulation rate (syll/s)	HNR (dB)	SE (dB)	LTAS slope (dB/100 Hz)	Pleasant
F1	237	−4.5	616	2.5	6.1	12.2	1.6	−20.6	3
F2	182	−2.0	683	2.1	7.2	11.0	1.9	−16.0	4
F3	212	−3.0	553	2.6	5.1	13.4	4.9	−14.7	3
F4	177	−2.7	670	2.5	6.5	12.7	1.6	−15.0	3
F5	212	−4.5	494	2.3	5.5	8.8	4.4	−9.8	2
M1	115	−2.4	761	2.7	6.1	7.5	1.6	−18.6	4
M2	107	−1.5	603	2.1	6.1	7.4	2.6	−14.1	4
M3	141	−2.8	555	2.8	5.7	6.7	4.8	−6.8	3
M4	175	−3.6	441	2.1	5.5	12.5	1.4	−19.0	3
M5	127	−3.3	652	2.4	5.5	7.5	2.4	−16.9	5
IL	101	−1.9	833	2.5	5.8	6.5	2.3	−19.9	5

IL is the Brazilian actor.

the highest pause duration mean, and one of the fastest LTAS slopes.

According to the LDA model, it can be seen that higher articulation rates and lower spectral emphasis tend to have higher scores of chunk-size pleasantness (final column).

For the sake of examining the consequences for the perception of less-known prosodic parameters, **Figure 6** presents an illustration of the F0 contours for the reading of a one-verse chunk by two Brazilian female speakers. The reading corresponding to the contour with sharper decreasing slopes

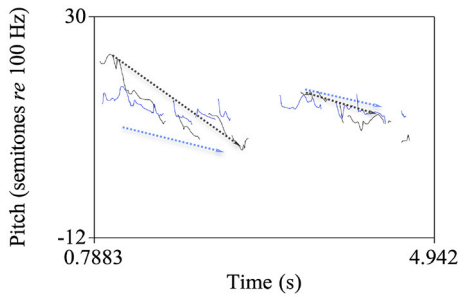


FIGURE 6 | F0 contours of the reading of a verse by two Brazilian female speakers in semitones ref to 100 Hz. The black one, with faster falling slopes, was considered less pleasant. The dotted arrows help signal the slopes of the falls for the sake of helping the reader.

from the F0 peaks (black) was considered less pleasant. The direction of this choice happened 61% out of all choices.

If the direction of the differences in the degree of pleasantness attributed to the whole poem's reciter by the respective linguistic community is used as a prediction of the response in the discrimination test, the significant correlations between predicted-from-Likert-test and actual discrimination tests are 54% for Brazilian listeners and 63% for Portuguese listeners. For that correlation, one of the vectors was formed by the difference of the median degrees from the Likert-scale test considering the recitation from each chunk was extracted for building the pair and, the other vector, the responses of the discrimination test: 2, if the second reading is more pleasant than the first and 1, in the opposite case.

DISCUSSION

The most relevant parameters for predicting pleasantness in the evaluation of both the whole poem and the comparison of the respective chunk readings point to the classes of pausing and tempo (pause duration, pausing rate, and articulation rate), voice quality (spectral emphasis and LTAS slope) and melody (F0 falls), with a minor prediction power for F0 median in Brazilian and Portuguese male reciters, according to the models summarized in **Table 1**. On the other hand, the degree of pleasantness attributed to the recitation of the whole poem by Portuguese female reciters depends only on pause duration. It is also important to signal a word of caution as regards the models involving the evaluation of wellbeing caused by female recitations in both varieties of Portuguese due to the extremely low explained variance, only 5%, despite the significance of the predictors in the final models. Because the correlation of that evaluation with the one of pleasantness is superior to 75%, the reasons for the weak correspondence with acoustic parameters should be looked at something non-acoustic in the female recitations, such as the relation of some parameters with the content, for instance, a pause realized in a particular word boundary. This can only be disentangled by coding these aspects and including other poems in future research.

The prediction of wellbeing degree in the Likert-scale test is main explained by parameters related to pausing (duration and rate), voice quality (LTAS slope and HNR), and F0 median as far as Brazilian reciters are concerned. The evaluation of the wellbeing caused by Portuguese recitations concerns the same domains but is restricted to the parameters of pause duration, spectral emphasis, and F0 median. For this variety, the results of the evaluation of male recitations are more relevant, due to higher explained variance.

Actors' recitations were preferred, both in direct comparisons and in the whole poem recitation, where pause rate and duration are highly relevant, parameters for which the actors exhibit the longest median pauses. The Portuguese actor also has the lowest rate of pause, the highest articulation rate, and the smoothest F0 mean falls.

It is important to point out, as far as pleasantness and wellbeing are concerned, that the following parameters were not significant predictors in any case: time position of F0 local peaks, F0 peak rate, F0 mean peak bandwidth, rates of F0 rise mean, the coefficient of variation of global intensity (CVInt), jitter, and shimmer. As we have seen in the introduction, jitter is associated with pleasantness in German, and its role in Portuguese is found in EP as a parameter for predicting pleasantness, as seen in the work by Pinto-Coelho et al. (2013) where low values of this parameter and for shimmer are preferred. As for the other parameters, their low variation across the chunks, especially CVInt (with an intra-speaker/inter-speaker ratio varying below 1%), indicates that a higher variation would be more related to liveliness than to the scales measured here. The theme of the poem, the insignificance of the poet in comparison with the passage of the seasons, especially the return of Spring, does not favor a lively recitation.

As for the hypotheses raised in the Introduction, we have already seen that Hypothesis 1 (pleasantness and wellbeing appraisals are equivalent, but the second has a weaker relation to acoustics because it is more subjective than the first) and Hypothesis 11 (Actors' voices are more pleasant and cause more wellbeing than those of lay reciters) were confirmed. In the following, the other hypotheses will be examined.

Hypotheses 2 (lower mean F0 increases both the sensations of pleasantness and wellbeing for women listening to the recitation of men) and 3 (higher mean F0 increases both the sensations of pleasantness and wellbeing for men listening to the recitation of women) were not entirely confirmed because F0 median was significant only in the evaluation of male recitations for which higher-pitched men were preferred by the two Brazilian genders of listeners, whereas lower-pitched men were preferred by the two Portuguese genders of listeners. Because articulation rate was significant only in the case of paired chunks and in the direction of higher rates being preferred and the rate of pauses predicted less frequent pauses as being more pleasant, Hypothesis 4 (slowing down and more frequent pauses increase both pleasantness and wellbeing) was not confirmed. Hypothesis 5 (breathier speech increases pleasantness and wellbeing) was not confirmed because higher values HNR increases the sensation of wellbeing and not lower values, which are related to more roughness and breathiness. This is in relation to the work by

Madureira (2008) for BP, where the speaker with a whispery voice sounded sad. Hypothesis 6 (more variable F0 increases pleasantness and wellbeing) was not confirmed because F0SAQ was not a significant predictor. Hypothesis 7 (the lesser vocal effort, the higher the sensations of pleasantness and wellbeing is) is confirmed because the spectral emphasis was significant for Portuguese male reciters for both pleasantness and wellbeing, for which the lower values increase both sensations. The high correlations of responses between Brazilian and Portuguese listeners broke down according to the variety of the reciter disconfirms Hypothesis 8 (Portuguese listeners appreciate the declamation by Brazilians more than Brazilians appreciate the declamation by Portuguese speakers).

Hypotheses 9 and 10 were not confirmed due to the differences in the models. As for Hypothesis 9 (the acoustic parameters related to pleasantness and wellbeing are the same and in the same degree of relative importance in the two varieties of Portuguese), although there is a close correspondence between the acoustics of the reciters' varieties of Portuguese, meaning that rhythmic parameters such as pause duration and articulation rate are significant predictors, with F0 median added in the case of male reciters in both varieties and measures of vocal effort, namely spectral emphasis and LTAS slope, the order of importance and specific combination of predictors is not the same in the two varieties of Portuguese. And because the predictors change depending on the gender of the reciter, Hypothesis 10 (men and women use the same prosodic-acoustic parameters to make the speech better appreciated by the listener) is not confirmed.

One key component of the acoustics of pleasantness, in particular in the models with higher amounts of variance explained, seems to be pausing because recitations with longer pause durations and lower pausing rates are more pleasant when our participants listen to the whole poem declamation. In pair comparisons, higher articulation rates, another temporal parameter, got higher degrees of pleasantness, which is in accordance with the higher attractiveness of faster male speech judged by women in the literature reviewed in the Introduction. Our results show that, at least for pleasantness, higher articulation rates are preferred by both male and female listeners irrespective of the gender of the reciter.

As far as the training of professional speakers is concerned, our results suggest the following relation between the use of voice/speech and the increase of the aesthetical pleasure in poetry declamation as related to sound pleasant and causing an increase of wellbeing in the listener: long pauses, more frequent pauses, higher articulation rates, lower vocal effort (related to sharper LTAS slopes and lower values of spectral emphasis).

CONCLUSIONS

The acoustics of pleasantness has some points in common with the acoustics of attractiveness when our results are compared with the findings reviewed in the Introduction mainly for English and German. One of these points is the role of tempo and vocal effort, as highlighted in the previous section. Despite minor differences in the evaluation of pleasantness depending on the variety of the reciter, the overall picture is quite the same, with the main role for tempo and voice quality through vocal effort acoustic correlates. The role of F0 falls in pair comparisons must be further investigated, but could also be related to slowing down, because smoother (slower) falls were considered more pleasant.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found at: https://figshare.com/authors/Plinio_Barbosa/11320902.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Comitê de Ética em Pesquisa from the University of Campinas. The participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and has approved it for publication.

FUNDING

The author thank our reciters and listeners from the two sides of the Atlantic Ocean and the CNPq grant # 302194/2019-3. The author also in debt with research assistant Alice Crochiquia for her help with the online perception tests. This work was registered in the Ethics Committee under number 41315420.8.0000.8142.

REFERENCES

Anderson, R. C., Klofstad, C. A., Mayew, W. J., and Venkatachalam, M. (2014). Vocal fry may undermine the success of young women in the labor market. *PLoS ONE* 9, e97506. doi: 10.1371/journal.pone.0097506

Barbosa, P. A. (2020). *Prosody Descriptor Extractor*. Available online at: https://github.com/pabarbosa/prosody-scripts/tree/master/ProsodyDescriptorExtractor.

Barbosa, P. A., and Eriksson, A., Åkesson, J. (2013). "Cross-linguistic similarities and differences of lexical stress realisation in Swedish and Brazilian Portuguese," in *Nordic Prosody XI, 2012 Tartu Nordic Prosody; Proceedings From the XIth Conference*. Frankfurt am Main: Peter Lang.

Baumann, T. (2017). "Large-scale speaker ranking from crowdsourced pairwise listener ratings," in *Proceedings of INTERSPEECH 2017*. (Stockholm), 2262–2266. doi: 10.21437/Interspeech.2017-1697

Boersma, P., and Weenink, D. (2021). *Praat: Doing Phonetics by Computer*. Available online at: http://www.praat.org/

Braga, D., Coelho, L., Resende, F. G. V. Jr., and Dias, M. S. (2007). "Subjective and objective evaluation of Brazilian Portuguese TTS voice font quality,"

in *Advances in Speech Technology, 14th International Workshop, June 27–29 2007* (Maribor).

Burkhardt, F., Eckert, M., Johannsen, W., and Stegmann, J. (2010). "A database of age and gender annotated telephone speech," in *Proceedings of LREC 2010* (Valletta), 1562–1565.

Camargo, Z., Madureira, S., Reis, N., and Rilliard, A. (2019). "The phonetic approach of voice qualities: challenges in corresponding perceptual to acoustic descriptions," in *Subsidia: Tools and Resources for Speech Sciences*, eds J. M. Lahoz-Bengoechea, and R. P. RamÃşn (Malaga: University of Malaga), 11–17.

Carlsen, J. M., Franke, M., Huttner, L. M., and Radtke, A. (2018). "How to measure a pleasant voice," in *Proceedings of the Conference Phonetics and Phonology in the German Language Area (P&P14)* (Vienna), 22–26.

de Krom, G. (1994). "Spectral correlates of breathiness and roughness for different types of vowel fragments," in *Proceeding 3rd International Conference on Spoken Language Processing (ICSLP 1994).* (Yokohama), 1471–1474.

Eriksson, A., Barbosa, P. A., and Åkesson, J. (2013). "The acoustics of word stress in Swedish: a function of stress level, speaking style and word accent," in *Proceedings of the 14th Annual Conference of the International Speech Communication Association, Lyon (Interspeech 2013)* (London: Causal Productions), 778–781. doi: 10.21437/Interspeech.2013-226

Eskénazi, M. (1993). "Trends in speaking styles research," in *Proceedings of the Third Eurospeech* (Berlin), 501–509.

Fónagy, I. (1961). Communication in poetry. *Word* 17, 194–218. doi: 10.1080/00437956.1961.11659754

Fónagy, I. (2001). *Languages Within Language: An Evolutive Approach.* Philadephia, PA: John Benjamins Publishing. doi: 10.1075/fos.13

Hodges-Simeon, C. R., Gaulin, S. J., and Puts, D. A. (2010). Different vocal parameters predict perceptions of dominance and attractiveness. *Hum. Nat.* 21, 406–427. doi: 10.1007/s12110-010-9101-5

Jacobsen, T. (2010). Beauty and the brain: culture, history and individual differences in aesthetic appreciation. *J. Anat.* 216, 184–191. doi: 10.1111/j.1469-7580.2009.01164.x

Khan, S. U. D., Becker, K., and Zimman, L. (2015). Acoustic correlates of creaky voice in English. *J. Acoust. Soc. Am.* 137, 2267–2267. doi: 10.1121/1.4920276

Laver, J. (1980). *The Phonetic Description of Voice Quality*. Cambridge: Cambridge University Press.

Léon, P. (1993). *Précis de Phonostylistique.* Paris: Nathan, p. 157–184.

Löfqvist, A., and Mandersson, B. (1987). Long-Time average spectrum of speech and voice analysis. *Folia Phoniatr. Logopaed.* 39, 221–229. doi: 10.1159/000265863

Madureira, S. (2008). "Reciting a sonnet: production strategies and perceptual effects," in *Proceeding Speech Prosody 2008*. (Campinas), 697–700.

Madureira, S., and Camargo, Z. (2010). "Exploring sound symbolism in the investigation of speech expressivity," in *Proceedings of the 3rd ISCA Tutorial and Research Workshop on Experimental Linguistics (ExLing 2010),* ed A. Botinis (Athens: ISCA and University of Athens), 105–108.

Madureira, S., and Fontes, M. A. S. (2019). "The analysis of facial and speech expressivity: tools and methods," in *Subsidia: Tools and Resources for Speech Sciences*, eds J. M. Lahoz-Bengoechea, R. P. RamÃşn (Malaga: University of Malaga), 19–26.

Master, S., Biase, N. D., Pedrosa, V., and Chiari, B. M. (2006). The long-term average spectrum in research and in the clinical practice of speech therapists. *Pró Fono Rev. Atualiz. Científ.* 18, 111–120. doi: 10.1590/S0104-56872006000100013

Mendoza, E., Valencia, N., Muñoz, J., and Trujillo, H. (1996). Differences in voice quality between men and women: use of the long-term average spectrum (LTAS). *J. Voice* 10, 59–66. doi: 10.1016/S0892-1997(96)80019-1

Menegon, P. S., Madureira, S., and Fontes, M. (2021). Análise de interpretações orais de um poema quanto aos aspectos expressivos. *Intercâmbio* 49, 113–137. doi: 10.23925/2237.759X.2021V49.56747

Niebuhr, O., Thumm, J., and Michalsky, J. (2018). "Shapes and timing in charismatic speech–Evidence from sounds and melodies," in *Proceeding 9th International Conference of Speech Prosody* (Poznań). doi: 10.21437/SpeechProsody.2018-118

Pinto-Coelho, L., Braga, D., Sales-Dias, M., and Garcia-Mateo, C. (2013). On the development of an automatic voice pleasantness classification and intensity estimation system. *Comp. Speech Lang.* 27, 75–88. doi: 10.1016/j.csl.2012.01.006

Quené, H. (2020). "Attractiveness of male speakers: effects of pitch and tempo," in *Voice Attractiveness: Studies on Sexy, Likable, and Charismatic Speaker*, eds B. B. Weiss, J. Trouvain, M. Barkat-Defradas, and J. J. Ohala (Singapore: Springer), 153–164. doi: 10.1007/978-981-15-6627-1_9

R Development Core Team (2008). *R: A Language and Environment for Statistical Computing.* Vienna: R Foundation for Statistical Computing. Available online at: http://www.r-project.org

Roback, P., and Legler, J. (2021). *Beyond Multiple Linear Regression: Applied Generalized Linear Models and Multilevel Models in R.* London: Chapman and Hall/CRC. doi: 10.1201/9780429066665

Rosenberg, A., and Hirschberg, J. (2020). "Prosodic aspects of the attractive voice," in *Voice Attractiveness: Studies on Sexy, Likable, and Charismatic Speaker*, eds B. B. Weiss, J. Trouvain, M. Barkat-Defradas, and J. J. Ohala (Singapore: Springer), 17–40. doi: 10.1007/978-981-15-6627-1_2

Severin, F., Bozkurt, B., and Dutoit, T. (2005). "HNR extraction in voiced speech, oriented towards voice quality analysis," in *2005 13th European Signal Processing Conference* (Antalya: IEEE).

Stathopoulos, E. T., Huber, J. E., and Sussman, J. E. (2011). Changes in acoustic characteristics of the voice across the life span: measures from individuals 4–93 years of age. *J. Speech Lang. Hear. Res.* 54, 1011–1021. doi: 10.1044/1092-4388(2010/10-0036)

Strangert, E., and Gustafson, J. (2008). "What makes a good speaker? Subject ratings, acoustic measurements and perceptual evaluations," in *Proceedings of INTERSPEECH 2008*, 1688–1691. doi: 10.21437/Interspeech.2008-368

Sundberg, J. (1994). Perceptual aspects of singing. *J. Voice* 8, 106–122. doi: 10.1016/S0892-1997(05)80303-0

Tegnér, I., Fox, J., Philipp, R., and Thorne, P. (2009). Evaluating the use of poetry to improve well-being and emotional resilience in cancer patients. *J. Poetry Ther.* 22, 121–131. doi: 10.1080/08893670903198383

Traunmüller, H., and Eriksson, A. (2000). Acoustic effects of variation in vocal effort by men, women, and children. *J. Acoust. Soc. Am.* 107, 3438–3451. doi: 10.1121/1.429414

Verdonck-de Leeuw, I. M., Festen, J. M., and Mahieu, H. F. (2001). Deviant vocal fold vibration as observed during videokymography: the effect on voice quality. *J. Voice* 15, 313–322. doi: 10.1016/S0892-1997(01)00033-9

Weiss, B., and Burkhardt, F. (2010). "Voice attributes affecting likability perception," in *Proceedings of INTERSPEECH 2010* (Chiba), 2014–2017. doi: 10.21437/Interspeech.2010-570

Weiss, B., and Burkhardt, F. (2012). "Is 'not bad' good enough? Aspects of unknown voices likability," *in Proceedings of INTERSPEECH 2012*, 510–513. doi: 10.21437/Interspeech.2012-97

Weiss, B., Trouvain, J., and Burkhardt, F. (2020). "Acoustic correlates of likable speakers in the NSC database," in *Voice Attractiveness: Studies on Sexy, Likable, and Charismatic Speaker*, eds B. B. Weiss, J. Trouvain, M. Barkat-Defradas, and J. J. Ohala (Singapore: Springer), 245–262. doi: 10.1007/978-981-15-6627-1_13

Conflict of Interest: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Barbosa. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers | Frontiers in Psychology

ORIGINAL RESEARCH
published: 11 May 2022
doi: 10.3389/fpsyg.2022.837752

Effects of Social Media Usage on Consumers' Purchase Intention in Social Commerce: A Cross-Cultural Empirical Analysis

Shangui Hu[1]* *and Zhen Zhu*[2]

[1] College of Business Administration, Ningbo University of Finance and Economics, Ningbo, China, [2] School of Economics and Management, China University of Geosciences, Wuhan, China

OPEN ACCESS

Edited by:
Ellen Anke Schmid,
Munich University of the Federal Armed Forces, Germany

Reviewed by:
Eduardo Moraes Sarmento,
Lusophone University of Humanities and Technologies, Portugal
Muhammad Rehan Masoom,
United International University, Bangladesh
Taqwa Hariguna,
Amikom University Purwokerto, Indonesia

***Correspondence:**
Shangui Hu
stanleyhff@hotmail.com

Specialty section:
This article was submitted to Human-Media Interaction, a section of the journal Frontiers in Psychology

Received: 17 December 2021
Accepted: 08 March 2022
Published: 11 May 2022

Citation:
Hu S and Zhu Z (2022) Effects of Social Media Usage on Consumers' Purchase Intention in Social Commerce: A Cross-Cultural Empirical Analysis. Front. Psychol. 13:837752. doi: 10.3389/fpsyg.2022.837752

Social commerce has produced enormous economic benefits as well as challenges for organizations, individuals, and industries. However, social media usage does not necessarily generate users' intention to purchase on social commerce websites. How social media usage influences users' purchase intention on social commerce websites still deserves more scholarly attention and this seems particularly important when social commerce transcends borders and countries. Taking an interdisciplinary perspective, the current study adopted a survey research method and identified the roles of social media usage in arousing users' purchase intention on social commerce websites in a culturally diversified environment. The data was collected from 2,058 international students coming from 135 countries and was analyzed using MPLUS based structural equation modeling. The research unveils the pathway whereby social media usage serves to generate users' purchase intention on social commerce websites. Importantly, users' cultural intelligence has been found to play a significant role mediating the effects of social media usage on users' intention. Further, cultural distance was found to attenuate the effects of social media usage on cultural intelligence. Based on the research findings, the study suggests that social commerce practitioners should be fully aware of the enabling roles of social media and cultural intelligence as well as the deterring role of cultural distance when arousing customers' purchasing intention in cross-cultural business operations. Any measures facilitated by social media usage to improve international consumers' cultural intelligence and mitigate the negative effects of cultural distance are supposed to be effective to enhance their purchasing intention. Accordingly, the study confirms the mutually melt and integrative relationships between information technology advancement and business prosperity in cross-cultural environment, which eventually contribute to sustainable development of society.

Keywords: social media usage, purchase intention, cultural intelligence, cultural distance, social commerce, cross-cultural background

INTRODUCTION

The advancements of Web 2.0 technologies and the burgeoning applications of social media have dramatically revolutionized the performing effectiveness of business structures, models and operations (Huang and Benyoucef, 2015; Busalim and Hussin, 2016; Hariguna and Rachmawati, 2019; Zheng et al., 2020; Algharabat and Rana, 2021; Adiandari, 2022). One notable evolution is the emerging phenomenon of social commerce, which profoundly changes the business operations from vendor-push models into a user-driven environment (Wigand et al., 2008; Curty and Zhang, 2013; Lu et al., 2016; Alnoor et al., 2022). As a new emergence of business operations, social commerce allows consumers to interact actively with their peers to share commercial-related information and influence other consumers' online purchase behaviors (Lin et al., 2017).

Although social commerce has achieved tremendous success since it was coined as a new form of business practices, it is reported that almost half of social media users showed no purchase intention on social commerce websites (Ko, 2018). More evidences have been found that people are not necessarily influenced by their peers using social media for online shopping behaviors because of potential negative effects including poor customer service and websites quality (Gutama et al., 2021), inaccurate information, privacy disclosure and fraud on consumers (Lickel et al., 2000; Mikalef et al., 2017; Paireekreng et al., 2019; Wang and Herrando, 2019; Chawla and Kumar, 2021). In other words, users' purchase intention is not equally influenced by their connections with other users shopping online, and the behavior of using social media does not necessarily increase the likelihood of users' purchase on social commerce websites (Kim and Park, 2013). As such, it is of great significance to examine the mechanism whereby individuals' usage of social media generates their purchase intention on social commerce sites.

To keep up with the globalization pace of the 21st century, successful business relies heavily on international market and world-wide business operations (Chen et al., 2010). Accordingly, social commerce has integrated into the world economy as an important component of business operations in multinational companies (MNCs) (Friedman, 2006). Take alibabagroup.com as an example, the company offers millions of products to international consumers from over 190 countries each year (Deng and Wang, 2016). This suggests that social commerce transcends national borders and rates itself as an alternative option for both local and international consumers (Silbiger et al., 2021). However, cultural features, which to a large extent differentiate one country from another, vary in countries no matter how closely those countries are connected (Frijns and Garel, 2021). In this regard, cultural differences between countries create a more complex international business environment for social commerce operation. Albeit that few studies have identified the effects of culture on customer's perceived benefits, word-of-mouth, service quality perceptions and business relationship quality (Schumann et al., 2010; Hoppner et al., 2015; Tang, 2017), how cultural differences influence consumers' engagement in social commerce needs more empirical studies. As previous research indicated, increasing consumers' cognitive, motivational and behavioral capabilities to deal with cross-cultural issues is an efficacious way to mitigate the negative consequences brought by cultural differences on their psychological and cognitive changes (Hu et al., 2020). However, there is paucity of literature explicating the role of cultural differences played in the cross-cultural business environment. In light of the globalized economy, social commerce has demonstrated its power to facilitate the rapid growth of cross-border commerce. As dominating features distinguishing one country from another, cultural differences deserve more scholarly efforts. In particular, there is lack of a complete picture about how cultural differences affect the process of developing consumers' emotional and cognitive capability by leveraging on social media, thereby influencing their purchase intention when they surf social commerce websites in an environment with foreign cultural novelty. Therefore, the current study endeavors to address the above mentioned research gaps by answering the following questions:

Question 1: *In social commerce operations, how different dimensions of social media usage affect consumer's purchase intention, particularly in cross-cultural business settings?*

Question 2: *How is the functioning of social media usage on consumer's purchase intention influenced by cultural differences?*

In order to address the above research questions, we adopted the perspective of the social learning theory (Bandura, 1978). Scholars argue that the learning approach changes consumers' cognitive and affective appraisal first before influencing their purchase intention (Chen et al., 2017; Friedrich et al., 2019). More specifically, we contend that social media usage, whether for seeking information or socializing, improves users' cognitive and emotional aspects through interaction among community members and information exchange in social networks (Huang and Benyoucef, 2013; Hu et al., 2021a). The learning behaviors contribute to consumers' knowledge and understandings of foreign cultures embedded in relevant products or conveyed through social commerce websites. During the process of users' interaction with the environment and interplay between cognitive and emotional functions during the learning process, social media usage changes their attitudes toward the products and social commerce websites, and eventually influences their purchase-making intention and their final behaviors (Lorenzo et al., 2012).

This research makes several important contributions to the extant literature about social commerce in cross-cultural business environment. First, whereas loads of research work has focused on the functionalities of social media usage and antecedents determining consumers' purchase intention (Fu et al., 2020; Jin et al., 2022), this study enriches the extant literature by examining the respective effects exerted by two dimensions of social media usage on consumers' purchase intention in social commerce operations. Second, the majority of relevant studies have examined the direct effects of website features of social media platforms on consumers' purchase intention. The study explicates the underlying mechanism whereby social media usage influences consumers' purchase intention in social commerce environment. Third, in terms of the complexity of different cultural values, this

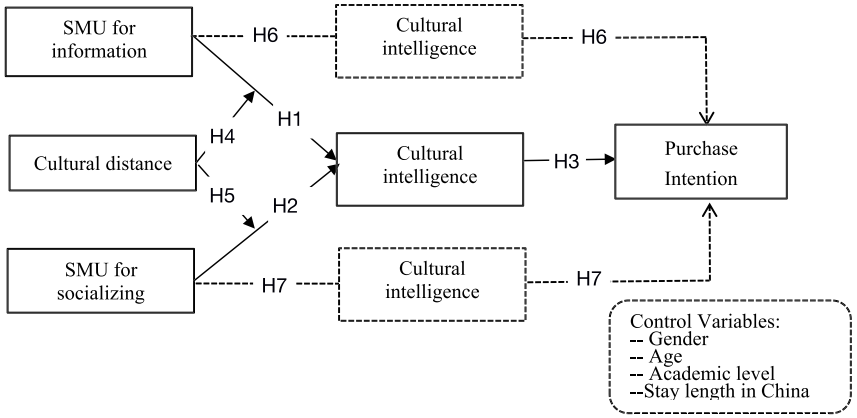


FIGURE 1 | Proposed research model. The dotted line refers to two mediation effects of cultural intelligence.

study further examines how the extent of cultural differences acts as a noisy channel to affect the functioning of social media usage on potential consumers' cultural intelligence, thereby influencing consumers' purchase intention on social commerce websites. Fourth, the study expands the extant research on social commerce into far more complicated cross-cultural settings. Given that cross-cultural business environment is characterized with more unexpected risks and challenges than a singular cultural context, the study explicates the significant role of cultural intelligence invigorating the success of social commerce in cross-cultural business backgrounds.

The remainder of the paper contains the following parts. The second section is about the literature review, which is followed by hypotheses development. Thereafter, we presented the research methodology and data analysis. Based on data analysis, discussion part was presented. After that, research limitations and avenues for future efforts are discussed. And the conclusion part ends the whole paper. The proposed research model was depicted in **Figure 1**.

LITERATURE REVIEW AND PROPOSED HYPOTHESES

Social Commerce and Social Learning Theory

Defined as "exchange-related activities that occur in, or are influenced by, an individual's social network in computer-mediated social environments, where the activities correspond to the need recognition, pre-purchase, purchase, and post-purchase stages of a focal exchange" (Yadav et al., 2013, p. 312), the model of social commerce has been created as a result of the advancement of information technology and communication tools (Hajli, 2015; Sun et al., 2019; Nacar and Ozdemir, 2022). As a new emergence of business operation model, social commerce enables vendors to interact with consumers through social media platforms and enables consumers themselves to communicate and exchange information about vendors' offered services and products (Aral et al., 2013; Friedrich et al., 2019). Three major attributes have been identified for social commerce operations, namely information technologies, social interactions, and organized commercial activities (Liang and Turban, 2011; Wang C. et al., 2019). Mediated by social media, social commerce involves the convergence of both offline and online environments (Zhou et al., 2013; Jami Pour et al., 2022). Two distinct categories of social commerce are summarized according to different business practice modes: (a) commercial activities that happen in traditional e-commerce websites where social communication technologies have been added to facilitate interaction between consumers themselves and between retailers and consumers; (b) commercial activities that happen in social networking sites where commercial features including transactions and advertisements have been incorporated (Zhang and Benyoucef, 2016; Chen et al., 2017).

To meet up with the demands of a globalized economy, multinational companies cross industrial and national borders to seek, access, and acquire new avenues for benefits attainment (Gassmann et al., 2010). In recognition of the conveniences and low costs yielded by applications of social media, business practitioners start their investment to development the platforms where social commerce is grounded. In such cases, social commerce serves as an innovative channel to facilitate firms to navigate international markets and access both local and international consumers through internet connections (Kotabe et al., 2002). Although the social commerce is burgeoning worldwide, yet there are still tough encounters that impede consumers' purchase intention and shopping behaviors (Van Heel et al., 2011; Sun et al., 2019; Algharabat and Rana, 2021). Previous research has identified that e-commerce conducted internationally has to deal with more complex cross-cultural settings where cultural differences between countries exist as dominating stressors hampering consumers' online experiences (Hu et al., 2018, 2020). In this aspect, more studies conducted with empirical methods are complementary measures to unveil how social commerce consumers' intention to purchase is aroused in more complex settings featured by cultural novelties and uncertainties.

From the perspective of social media users, when they browse social media networks to search for relevant products' information and interact with others for exchanging ideas, they are in actuality learning from others before they develop their purchase intention (Huang and Benyoucef, 2013). From the perspective of social learning theory (SLT), all learning is underpinned with two fundamental assumptions. The first one is about the external interaction between learners and their exposed social, cultural, and material environments, and the second refers to the internal psychological process of knowledge and skills acquisition and explication based on previous learning. Only when the two processes are involved together, not separately, can learning behaviors take place. In parallel with two underlying processes through which learning takes place, three dimensions of learning have been depicted, namely cognitive dimension, emotional dimension, and social dimension (Illeris, 2003; Chen et al., 2017). The cognitive dimension is covered in the internal process and deals with the learning content about knowledge and skills which builds up learner's cognitive framework and their ability. Likewise, the emotional or psychodynamic dimension is included in the internal process and encompasses necessary mental energy, feelings, and motivations for completing the process. The social dimension of learning depicts the external interaction between learners and their environment including their participation, communication, cooperation, and even integration in communities and society (Illeris, 2003). It is noteworthy that the three dimensions of learning are interdependent and embedded in the social, cultural, and material environments, and the psychological process is crucially dependent on the external interactions. Since its initiation, SLT has demonstrated its explanatory power to explicate human's development and behaviors, and has been adopted to elaborate on customers' social commerce behaviors (Chavis, 2011; Hu et al., 2018).

The current research adopts the social learning theory as the theoretical lens because the theory offers a structured approach as a response to explicate an individual's psychological changes and behavioral patterns in different social and cultural contexts (Illeris, 2003; Chen et al., 2017). Albeit social commerce demonstrates itself as a special form of e-commerce, the fundamental social psychology backing up consumers' online shopping intention or behaviors still lies in social learning (Chen et al., 2017). SLT indicates that consumers' purchase-making decision or intention is determined both by an external interaction between consumers and their exposed environments, and by an internal psychological process to acquire and elaborate environmental elements. In a cross-cultural environment, social media platforms provide conditions for consumers to learn about foreign cultural knowledge posted on the websites or embedded in offered products (Hu et al., 2021a). Further, consumers also learn from observations of others' social commerce experiences by means of information exchange and interaction with community members (Hu et al., 2017). During the external interaction with environmental factors, social media users experience cognitive, emotional, and behavioral changes which were originally interpreted as their development of the capacity to manage cross-cultural encounters, namely cultural intelligence (Ang et al., 2007; Hu et al., 2018). According to SLT, those changes may finally arouse users' purchase-making decision or intention as a result of an integrated interplay between the functions of cognition and emotion during the process of learning.

Social Media Usage (Informational and Socializing) and Cultural Intelligence

Due to the advancement of information technology, social media has permeated into every territory of contemporary society and exerts vast impacts on individuals' and organizations' survival and success by functioning as a platform for individuals' social interaction and information exchanging (Kauffman et al., 2017; Mohammed and Qhal, 2020; Lin et al., 2021). Strong evidence indicates that individuals are driven by utilitarian and hedonic value orientations when they use social media platforms (Wang, 2010), and generally gratified with informational and socializing purposes (Hu et al., 2020). More specifically, social media usage for informational purpose is to gratify individual cognitive needs by searching for information to solve confronted problems, whilst social media usage with socializing motivation is to satisfy individual emotional needs by establishing and maintaining interpersonal relationships through networking sites (LaRose, 2009; Hu et al., 2021b; Guan et al., 2022).

When accommodated in a culturally novel environment filled with cultural conflicts and potential risks, individuals are supposed to manage cultural conflicts and unexpected situations (Ward et al., 2011), which may trigger individuals' negative emotions such as loneliness, depression, and psychological disorder (Hu et al., 2020). Social media provides opportunities for individuals to search for desired information and establish and maintain social networks (Hughes et al., 2012; Khaola et al., 2022). In this vein, social media is closely associated with individuals' international exposure by serving to develop their cultural intelligence and facilitate their working and living conveniences including online shopping experiences.

Coined by Earley and Ang (2003), the conceptualization of cultural intelligence (CQ) represents people's capacities to manage cross-cultural issues in a foreign culturally novel environment. Since its initiation, CQ is regarded as a key capacity enabling international expatriates to adjust themselves to a culturally new land, develop personal abilities and conduct successful overseas businesses (Imai and Gelfand, 2010; Chen et al., 2012; Kim and Dyne, 2012). Numerous studies have identified that individuals with a higher level of CQ demonstrate more desirable outcomes, such as leadership performance (Guang and Charoensukmongkol, 2022), self-efficacy (Hu et al., 2018; AlMazrouei and Zacca, 2021), cross-cultural adjustment (Hu et al., 2020), organizational embeddedness and knowledge sharing (Stoermer et al., 2021), and global leadership (Ang and Van Dyne, 2015). On the other hand, individuals with a lower level of CQ may experience unexpected problems in their learning and working environments, such as discrimination, loneliness and depression in cross-cultural settings (Baltacı, 2017; Karatas and Arpaci, 2021). In light of the significance of CQ, scholars endeavored a lot to explore the antecedents

influencing the development of individuals' CQ. Studies have found that both environmental elements and personal attributes affect the development of individuals' CQ, such as multicultural experiences and training programs (Hu et al., 2017), information technology (Hu et al., 2018), self-efficacy (Hu et al., 2018) and personality dimensions (Presbitero, 2018; Hu et al., 2020). Albeit research has confirmed that CQ functions as a mediating mechanism (Hu et al., 2020), how CQ mediates the relationships between dimensions of social media usage and customers' purchase intention in cross-cultural social commerce environment still lacks sufficient empirical evidence.

Regarded as being multidimensional, cultural intelligence comprises four dimensions which cover metacognitive, cognitive, motivational, and behavioral aspects of CQ (Earley and Ang, 2003; Stoermer et al., 2021). As a fundamental element, metacognitive CQ is defined as an individual's cognitive power to control consciousness to acquire and comprehend knowledge about a foreign culture and make necessary mindset adjustment to expected and unexpected foreign situations (Ang et al., 2007; Hu et al., 2018). Cognitive dimension of CQ pertains to the systematized knowledge learned about the foreign culture individuals are exposed to, which includes legislated norms, underlying values, long-term traditions and social practices (Triandis, 2006). With good mastering of this kind of intelligence, individuals are able to master know-how to distinguish cultural differences existing among different cultures (Hu et al., 2020). Likewise, motivational dimension of cultural intelligence demonstrates an individual's sufficient motivations to cope with foreign cultural related issues successfully (Hu et al., 2021b). Accordingly, behavioral CQ represents an individual's behavioral ability to respond appropriately to expected and unexpected encounters in a culturally diversified environment (Hu et al., 2018; Stoermer et al., 2021), which contributes to individuals' general adjustment, working performance and learning results in a cross-cultural environment.

According to SLT, learning happens through external interaction processes and internal psychological process with cognitive and emotional changes. Firstly, when situated in a culturally diversified business environment, social media users are inclined to use social media to acquire related information for capture and comprehension of foreign cultural knowledge that are originally different from that of their home countries (Hu et al., 2018). Further, SLT suggests that the observation of others' behaviors and experiences is also a basic learning approach (Chen et al., 2017). Users observe others' shopping experiences by reading their posted information on social media and learn about their shopping experiences and comments shared on the social commerce websites (Bugshan and Attar, 2020; Stoermer et al., 2021). Such experiences facilitate users to expand their knowledge infrastructure and self-consciousness with attained values, conventions, values, and cultural practices in international business settings, which contributes to the development of individuals' both metacognitive and cognitive dimensions of CQ. In addition, established metacognitive CQ and cognitive CQ facilitate individuals with sufficient interest and motivation to be engaged in complicated cross-cultural issues. Moreover, more acquaintance with foreign cultures embedded in the products and social commerce websites arouses individuals to have more interests and motivations to manage effectively their behaviors and decision-making patterns, which contributes to their development of individuals' motivational CQ. Accordingly, equipped with orientated self-consciousness, sufficient cultural knowledge, and strong intrinsic motivation, individuals are supposed to behave appropriately in a culturally diversified social commerce environment (Glass and Westmont, 2014; Guang and Charoensukmongkol, 2022). Based on the above statements, we make the hypothesis as follows:

H1: Social media usage for informational purpose has positive associations with cultural intelligence.

Social learning theory underlines that personal development and psychological changes of human beings are triggered by the interaction effects between environment's external elements and individual's inner psychological changes (Bandura, 1978). Within the theoretical framework of SLT, social media can be acknowledged well as an environmental factor to interact externally with learners (Illeris, 2003; Hossain et al., 2021). As a social dimension of SLT, learners are actively involved in establishing and maintaining interpersonal relationships with community members on social media websites (Weller, 2016; Matsuda, 2022). In particular, international users tend to be more engaged in social media networking when they are exposed to novel culture and are confronted with numerous uncertainties and risks (Zhang and Goodson, 2011; Karatas and Arpaci, 2021). As such, they utilize social media to socialize with people who are culturally different (Hu et al., 2018). During the interaction process, social media users leverage on online opportunities to exchange necessary information related to cross-cultural issues with their peers, seniors, and colleagues from diversified cultural backgrounds. Moreover, the attained information from online interaction seems much more reliable, accurate, and trustworthy than that sought directly from business websites (Hur et al., 2017; Soleimani, 2021). Therefore, social media usage offers smart channels for potential consumers to interact with others from various cultural backgrounds for information, knowledge sharing and idea exchange purposes (Ghahtarani et al., 2020; Jen and Lin, 2021). This conduit increases their knowledge infrastructure and boosts their consciousness and motivations to learn more about another foreign culture (AlMazrouei and Zacca, 2021). Additionally, social media usage contributes to the development of maintaining extant interpersonal connections and establishing new social connections as well, which enables users to develop appropriate behavior capabilities in a culturally novel environment (Bonebrake, 2002; Hu et al., 2020; Rakhmansyah et al., 2022). Based on the above statement, we can logically hypothesize that:

H2: Social media usage for socializing purpose has positive associations with cultural intelligence.

Cultural Intelligence and Purchase Intention

In an online business environment, one cause of customer uncertainty is asymmetric information overloaded on users' decision making (Hwang et al., 2014). Users with lower levels of CQ are easy to be harnessed by overloaded information and fail to differentiate useful information from the useless (Hu et al., 2018; Yazdanshenas, 2021). In contrast, higher levels of CQ facilitates users to sort out the useful from a chaotic reservoir of information, enabled by acquired knowledge about the foreign culture, tradition, values embedded in products and business environments (Hu et al., 2020). In this regard, CQ enables social media users to leverage on useful information both from social media and comments made by other consumers, thereby facilitating to develop their purchase intention on social commerce websites.

Empirical studies indicated that individuals with CQ demonstrate to be more culturally engaged with locals and more acceptable than individuals with relatively lower level of CQ (Earley and Ang, 2003; Imai and Gelfand, 2010). This can be interpreted in the way that individuals with CQ have more cultural consciousness and cognitive knowledge to develop and maintain trustworthy interpersonal relationships by interacting with those in the online community with attributes of their intrinsic motivation and behavior capability (Ng, 2013). In particular, sufficient communications facilitate users to establish trust-based integration into the online community, and allow users to understand, predict, and meet the group-normative expectations (Ng, 2013; Guillaume et al., 2014; Prabowo et al., 2021). Further, research has demonstrated that trust is regarded as a basic psychological element to motivate social media users to surf social websites and demonstrate other online behaviors in social commerce environment (Stewart, 2003; Liang and Turban, 2011; Leong et al., 2020). This proposition supports the notion that well-established social relationships with online community members relieve users' concerns of being defrauded by online uncertainties and risky events, and thus drives them to develop purchase intention on social websites more possibly (Wang X. et al., 2019).

Thus it can be logically hypothesized that:

H3: Cultural intelligence has positive associations with users' purchase intention.

The Moderating Effects of Cultural Distance

As a fundamental element distinguishing one group from the other groups, culture is regarded as a contextual variable which influences individuals' motivation, attitudes and behaviors (Chen et al., 2018). In the competitive global market, culture is confirmed to repeatedly affect consumer's cognitive patterns and behaviors (McCort and Malhotra, 1993; Schumann et al., 2010). Cultural distance, interpreted as the extent of cultural differences distinguishing host country from individual's home country (Shenkar, 2001; Silbiger et al., 2021), describes the disparities existing between countries in terms of development level, education, language, business, legislated systems, conventions, cultural values, and connections level (Froese and Peltokorpi, 2011; Liu et al., 2021).

Since its evolution from physical distance, cultural distance has received burgeoning attention from scholars engaged in the literature of cross-cultural studies and international business. Accordingly, the conceptualization has been examined as an influential factor to test the robust effects of culture on international business patterns, trade flows, market entry, products' design philosophy and joint-venture performance (Froese and Peltokorpi, 2011; Silbiger et al., 2021). As one of the dominating stressors perceived by individuals in a culturally different community, cultural distance has been interpreted from the perspectives of cultural novelty and toughness (Mendenhall and Oddou, 1985; Black et al., 1991), which exerts effects on individuals' cross-cultural performance and psychological stability (Chen et al., 2010). However, there is a paucity of literature researching its effects on social commerce. Based on the social learning theory, scholars proposed that cultural distance is most likely to play a boundary conditional role influencing individual-level perceptions of stressors and stability of personal attributes, and thus brings negative effects to personal development, because it captures the most complicated aspects and challenges inherent in culturally diversified contexts (Chen et al., 2010).

When a high cultural distance between countries is perceived, it indicates that users have more inherent obstacles to comprehend the information when they use social media for informational purposes. In this regard, individuals feel more difficult to understand and accept cultural nuances and concepts embedded in products and social commerce websites. For the purposes of knowledge access and acquisition, when exposed to a foreign country with more cultural distance, users are supposed to allocate more personal resources to develop cultural intelligence by searching relevant information from social media. More specifically, when exposed to a culturally distant environment, users feel more inherent difficulties to search for, comprehend, and absorb appropriate information from social media for personal development. This is interpreted by the resource allocation model which illustrates that individuals are inhibited by lack of knowledge to manage assigned tasks despite of enormous efforts made (Yeo and Neal, 2004; Liu et al., 2021). Thus cultural distance is more likely to undermine the effects exerted by informational social media usage on individual's cultural intelligence.

H4: The effects of social media usage for information on individual's cultural intelligence are attenuated by cultural distance.

Likewise, an environment with higher cultural distance becomes more detrimental for users to socialize with others than the environment with lower cultural distance (Tett and Burnett, 2003; English et al., 2021). Previous studies have examined that cultural distance strongly impacts social relationships and personal communications in cross-cultural contexts (Galchenko and van de Vijver, 2007). Scholars

researching conflict management and cross-cultural studies consider that people from different cultures may find it more challenging to interact with each other (Ma, 2010). Accordingly, when users use social media websites with a socializing purpose to establish or maintain interpersonal relationships, they are deeply influenced by the homophily, which refers to the degree of cultural distance (Hu et al., 2020). A culturally distant business environment hinders users to adopt an effective strategy to communicate with those who have different conceptual understandings about culture and culturally rooted products, which finally minimize the effectiveness of using social media to socialize with others for developing individuals' cultural intelligence (Silbiger et al., 2021). Further, cultural distance impedes users from using social media websites to integrate into the social commerce community. For example, Chinese culture is relatively distant for users who are from western countries, because China is a typically eastern country which is deeply rooted in distinct collectivism culture, while western countries unanimously keep the value of individualism culture. As such, western consumers may fail to capture the integration of traditional cultural features reflected into Chinese products and business promotion strategies (Chen and Peng, 2008; Liu et al., 2021). Thus obstacles generated by cultural distance may result in failure of communication efforts and idea exchanges during the process of individuals' interacting with the local people through social media. In this regard, cultural distance separates users to be fully immersed into the online community. This also indicates that cultural distance prevents users from developing cultural intelligence when they use social media with the purpose of socializing with others in culturally novel settings. Thus it can be logically hypothesized that:

H5: The effects of social media usage for socializing on individual's cultural intelligence are attenuated by cultural distance.

The Mediating Effects of Cultural Intelligence

Scholars argue that social commerce is a multidimensional concept in actuality and consumers expected benefits from using social commerce come from one of three aspects or all of them, namely utility benefits, social benefits, and hedonic benefits (Osatuyi et al., 2020). Further, it has been identified that relationships between antecedents and users' intention to use social commerce could be non-linear, not direct, due to environmental and users' psychological factors' changes (Lankton et al., 2016; Osatuyi et al., 2020). As mentioned above, in cross-cultural contexts, the relationships between customers' social media usage and their purchase intention in social commerce deems to be more complicated in light of environmental complexities (Hu et al., 2018). In this regard, customers' potential interest to use social commerce relies heavily on seeking information and socializing with platform members they are familiar with or not through social media usage. As a result, potential customers' leveraging on social media for utilitarian and hedonic purposes result in changes of their cultural intelligence by increasing cognitive, motivational and behavioral capabilities. And those improved capabilities mobilize themselves to have better understandings about social commerce, thereby increasing their intention to use social commerce.

Since its creation, cultural intelligence has attracted scholars' attention due to its explanation power to elicit individuals' capabilities to deal with cross-cultural related issues (Ang et al., 2007; Pidduck et al., 2022). It is worthy to note that increasing research has uncovered that cultural intelligence plays as a key antecedent to predict individuals' positive psychological changes and cross-cultural behaviors (Ang and Van Dyne, 2015; Hu et al., 2020, 2021a). More importantly, cultural intelligence has been identified as a mechanism to explicate how and why individuals take actions and perform effectively in complicated cross-cultural environment (Lee and Sukoco, 2010; Hu et al., 2017; English et al., 2021). Accordingly, we contend that cultural intelligence acts as a mechanism as well when users utilize social media platforms to increase their purchase intention in social commerce by developing cultural intelligence as a prerequisite to better capture the content, design, delivered information and word-of-mouth about the products originated from different cultures. Numerous studies have also elicited that cultural intelligence plays the mediating roles pertaining to the final cross-cultural performances (Lee and Sukoco, 2010; MacNab and Worthley, 2012; Hu et al., 2021a).

Based on above statements, we hypothesize that:

H6: Cultural intelligence mediates the relationship between social media usage for information and purchase intention.

H7: Cultural intelligence mediates the relationship between social media usage for socializing and purchase intention.

METHODOLOGY

The study employed both Spss23 and Mplus8 statistical software to analyze the data. Descriptive statistics, means, variances and correlations of variables were calculated through Spss23. In addition, considering the advantages of Mplus8, this study carried out regression analysis by constructing structural equation models using Mplus8. Up to date, Mplus8 is regarded as one of the most popular and powerful software available for latent variable analysis. Firstly, Mplus8 can handle not only categorical latent variables but also continuous latent variable models as well as complex models with both variables. Moreover, it is specially designed for the analysis of various latent variable models, for missing values and for complex survey data. Secondly, Mplus8 can also analyze cross-sectional and longitudinal data, hierarchical data (multilevel analysis) and data from different aggregates (multi-group analysis). Finally, Mplus8 also offers a graphical manipulation function that allows the software to automatically help program and produce the corresponding results through graphical manipulation.

Sample and Procedure

To further examine the above proposed hypotheses, a survey was designed to measure the variables contained in the proposed

research model. With generous assistance from more than 40 public universities offering international education in China, the questionnaires were delivered to international students who were studying their degree or non-degree programs at those higher institutions. International students studying in China were chosen for the current research samples mainly for three reasons. First, international students are a special group of international consumers. Application of social media for information and social interactions has become a ubiquitous phenomenon among them. However, this special group of consumers is rarely investigated by scholars of social commerce research. Second, international students share similarities a lot with other international consumers confronted with unexpected risks and uncertainties in a multicultural environment (Hu et al., 2018). Third, most of the extant research was conducted in western cultural background. Meantime, China has made itself successfully be the first targeted country for international education in Asia with 492,185 international students registered in Ministry of Education of People's Republic of China (2018). Examining how social commerce is conducted out of western cultures is a necessary supplementation to the literature of social commerce research.

Before the large scale delivery of the questionnaires, two pilot tests were conducted to ensure the validity of all variables. For the first pilot test, 40 international students were invited to offer their responses to variable items. And 3 scholars engaged in relevant research fields were invited to assess the students' responses and variables' validity and make necessary revisions to the wording and paraphrases. Further, interviews were conducted among the first group pilot students to revise those items in order to ensure all those items to match their real *status quo* in China. To guarantee all revisions appropriate, a second pilot test was conducted among another group of 71 students after two weeks. The second pilot test demonstrated very good results. Afterward, the questionnaires were distributed through one of the largest survey platform Wenjuanxing[1], which is more efficiently used by scholars to collect sensitive data than offline channels (Stewart and Bing, 2009). The website links were offered to all volunteer students with the help of university officials. During the 1-month process of collecting data, reminded information was sent to students to increase the response rate. And a final of 2058 questionnaires from 135 countries' students were returned for data analysis. Among all participants, 1154 were male students and 904 were girls. Moreover, 65.89% students were from Asian countries and African students occupied 24.73%. Pertaining to their academic levels, more than 70% were bachelor or below bachelor degree holders. Further, almost half of all participants (46.65%) have been staying in China for more than 2 years. **Table 1** shows the demographic statistics of the sample.

TABLE 1 | Demographic statistics of the sample (N = 2058).

Category	Item	Frequency	Percentage (%)
Age	≤20 years old	365	17.74%
	21–25 years old	863	41.93%
	26–30 years old	473	22.98%
	≥31 years old	357	17.35%
Gender	Male	1154	56.07%
	Female	904	43.93%
Areas	Asia	1356	65.89%
	Africa	509	24.73%
	Europe	97	4.71%
	Oceania	17	0.83%
	America	75	3.64%
	Missing data	4	0.19%
Highest diploma	below bachelor	772	37.51%
	bachelor	681	33.09%
	master	493	23.96%
	doctor	112	5.44%
Stay length in China	Less than 6 months	446	21.67%
	6–12 months	127	6.17%
	1–2 years	525	25.51%
	More than 2 years	960	46.65%

Measurements

Three international students were chosen for the pilot test, who were participating in degree programs and language training courses, to complete the questionnaires to help evaluate the measures we adopted from previous literature. Necessary revisions were made to the question wordings according to three students' feedback to ensure the questionnaire items meet international students' real status in China. In addition, three academic researchers in the relevant fields were invited to join in the work of reviewing the validity of all variables. All the measurement work was conducted with five-point scales from 1 to 5. 1 represents users' "strongly disagree" and 5 represents users' "strongly agree."

For the measurement of social media usage, six items were adopted from Hughes et al. (2012). 6 items were divided into two halves. The first three items were used to measure social media usage for informational motivation and the other half were used to measure social media usage for socializing motivation.

To measure the overall cultural intelligence, we used 20 items, which were developed originally by Ang et al. (2007). To measure the metacognitive, cognitive, motivational and behavioral aspects, 4, 6, 5 and 5 items were used respectively.

Further, we measured users' purchase intention on social commerce websites by adopting three items proposed by Shin (2013). To measure the cultural distance, we used 6 items developed by Chen et al. (2010), to measure the extent of cultural differences between China and students' home countries. Sample items included "How similar are religions and rituals in China similar to your own country?" and "How similar are values in China similar to your own country?" Measures were assessed ranging from 1 for "very similar" to 5 for "not similar at all."

We also controlled for participants' age, gender, length of stay in China, and their education level to avoid any possible confounding effects identified in previous literature (Hu et al., 2020).

[1] www.wjx.cn

TABLE 2 | Means, standard deviations, and correlation matrix (N = 2058).

Constructs	Means	SD	CQ	CD	SMUI	SMUS	PI
CQ	3.9223	0.65314	**0.753**				
CD	3.3768	0.90273	0.307**	**0.825**			
SMUI	3.8431	0.90465	0.389**	0.342**	**0.953**		
SMUS	3.8481	0.86333	0.435**	0.324**	0.641**	**0.830**	
PI	3.9342	0.80811	0.568**	0.282**	0.372**	0.413**	**0.910**

*CQ, cultural intelligence; CD, cultural distance; SMUI, social media usage for information; SMUS, social media usage for socializing; PI, purchase intention; **$p < 0.01$ (two-tailed) and the diagonal is the arithmetic square root of the AVE value (bold values).*

TABLE 3 | Measurement of constructs.

Constructs	Dimensions	Loadings	CR	AVE	Cronbach's alpha
SMUI	SMUI1	0.853	0.766	0.908	0.843
	SMUI2	0.887			
	SMUI3	0.885			
SMUS	SMUS1	0.825	0.869	0.690	0.763
	SMUS2	0.791			
	SMUS3	0.873			
meta-cognitive CQ	CQ1	0.894	0.944	0.808	0.912
	CQ2	0.897			
	CQ3	0.916			
	CQ4	0.888			
Cognitive CQ	CQ5	0.823	0.932	0.696	0.912
	CQ6	0.810			
	CQ7	0.841			
	CQ8	0.845			
	CQ9	0.847			
	CQ10	0.840			
Motivational CQ	CQ11	0.833	0.931	0.728	0.970
	CQ12	0.847			
	CQ13	0.874			
	CQ14	0.852			
	CQ15	0.860			
Behavioral CQ	CQ16	0.776	0.928	0.723	0.903
	CQ17	0.848			
	CQ18	0.884			
	CQ19	0.885			
	CQ20	0.885			
PI	PI1	0.905	0.938	0.829	0.897
	PI2	0.912			
	PI3	0.915			
CD	CD1	0.792	0.929	0.688	0.906
	CD2	0.812			
	CD3	0.866			
	CD4	0.866			
	CD5	0.886			
	CD6	0.747			

CQ, cultural intelligence; CD, cultural distance; SMUI, social media usage for information; SMUS, social media usage for socializing; PI, purchase intention.

HYPOTHESES ANALYSIS

Constructs Validity

Preliminary analysis was performed to test potential common method bias, constructs' reliability and validity. The statistics shown in **Tables 2**, **3** show that the loadings for all standardized factors ranged from 0.747 to 0.915, which were above 0.6 criteria. The AVE values ranged from 0.688 to 0.908, greater than required 0.5 (Hair et al., 2020), and composite reliability was from 0.766 to 0.944, greater than required 0.7 (Bagozzi, 1981). All those statistics demonstrated the good convergent validity of all variables. Each variable's indicated arithmetic square root of the AVE value was greater than the correlation between this and other variables (Fornell and Larcker, 1981). Thus discriminant validity was confirmed.

Further, common method bias may exist when the response variations are the function of instrument rather than the real expressions of participants themselves. To assess the potential problem of common method bias, we used Harman's one factor, suggested by Podsakoff et al. (2012). Five principal components were created (with an eigenvalue greater than 1). The first principal component explains less than half of the total variance, which indicates that the common method bias is not a serious problem in the current research.

Hypothesis Testing

In order to test the above proposed hypotheses, we conducted a regression analysis with Mplus8.3. As mentioned above, we also controlled for participants' age, sex, educational level, and their stay length in China to avoid any potential confounding effects. The regression results were shown at **Figure 2**. Consistent with H1, informational usage of social media demonstrated positive relationship with cultural intelligence (r = 0.188, $p < 0.01$). Furthermore, socializing usage of social media demonstrated significantly positive relationship with cultural intelligence as well (r = 0.315, $p < 0.01$). Thus H2 was also supported. Thereafter, we also examined how cultural intelligence influences users' purchase intention. The result indicates that cultural intelligence also had positive association with users' purchase intention on social commerce websites (r = 0.462, $p < 0.01$). H3 was supported as well.

In addition, we examined the mediating functions of cultural intelligence in the proposed research model. A bootstrapping method was employed to test the mediation results. After 2000 iterations, cultural intelligence was found to mediate the relationships between two dimensions of social media usage and users' purchase intention respectively (SMUI-CQ-PI, 0.085, 95% CI [0.046, 0.124]; SMUS-CQ-PI, 0.145, 95% CI [0.105, 0.187]). The mediating effects were both significant for the 95% confidence interval. And there is no zero between the lower and the upper. Thus H6 and H7 were both supported. (see **Table 4**).

To examine the moderating effects of cultural distance, this study first decentralized the constructs of social media usage for information, social media usage for socializing and cultural distance. Following that step, we further created the interaction items between two dimensions of social media usage and cultural distance respectively and examined the effects of interactive

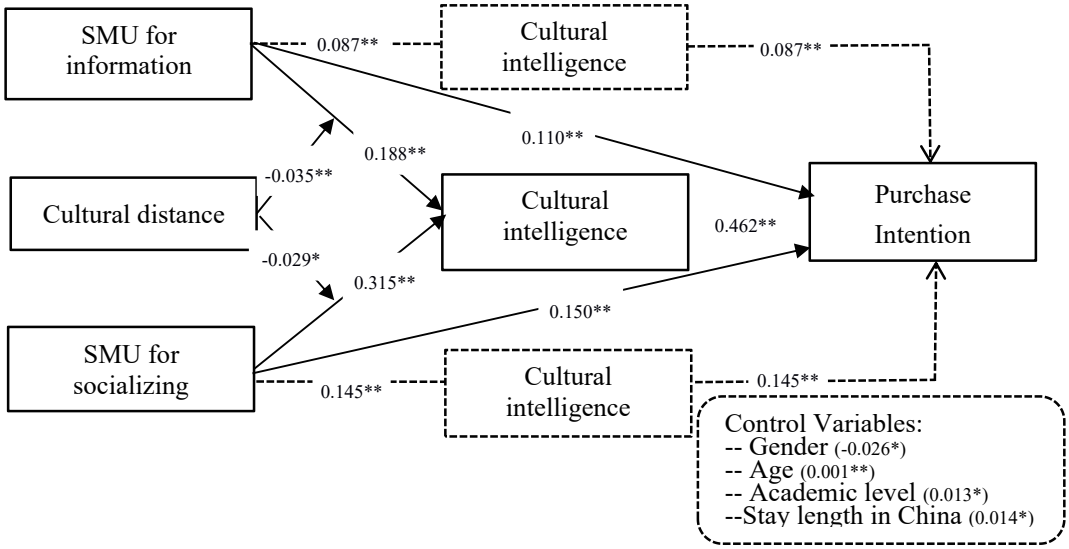


FIGURE 2 | Results of the hypothesized model. SMU, social media usage. *$p < 0.05$ and **$p < 0.01$.

TABLE 4 | Results of the mediating effect.

Mediating effect	Estimate	SE	*P*	BC 99% CI	
				Lower	Upper
SMUI-CQ-ICS	0.087	0.015	0.000	0.046	0.124
SMUS-CQ-ICS	0.145	0.016	0.000	0.105	0.187

TABLE 5 | Results of hierarchical regression analyses for interaction.

Variable	Cultural intelligence				
	Model 1	Model 2	Model 3	Model 4	Model 5
Age	–0.021	0.001	0.001	0.001	0.015
Gender	–0.003	–0.042	–0.044	–0.033	–0.034
Stay length in China	0.002	0.012	0.013	0.006	0.008
Highest diploma	–0.015	0.014	0.018	0.006	0.009
SMUI		0.237**	0.229**		
SMUS				0.285**	0.278**
CD		0.142**	0.145**	0.135**	0.137**
SMUI*CD			–0.035**		
SMUS*CD					–0.029*
R^2	0.001	0.189	0.192	0.221	0.223
ΔR^2		0.189	0.003	0.221	0.002
ΔF		79.593**	7.274**	97.235**	4.949**

*N = 2058. SMUI, social media usage for information; SMUS, social media usage for socializing; CD, cultural distance. Standardized regression coefficients are shown. *p < 0.05, **p < 0.01.*

constructs on cultural intelligence (**Table 5**). The result indicates that cultural distance attenuated the effects of informational social media usage on cultural intelligence ($r = –0.035$, $p < 0.01$). Thus H4 was supported. Further, cultural distance was also found to undermine the effects of socializing social media usage on cultural intelligence as well ($r = –0.029$, $p < 0.05$). Therefore H5 was supported as well.

The result of simple slope analysis further indicates that cultural distance mitigated the effects of social media usage, whether for information or socializing, on cultural intelligence (**Figures 3**, **4**).

DISCUSSION

The research adopted an empirical method and examined the associations among social media usage, users' purchase intention on social commerce websites, cultural distance and cultural intelligence. All proposed hypotheses were supported well after the regression analysis of the data collected from 2058 international students who came from 135 countries all over the world. The results indicate that two dimensions of social media usage, information and socializing, both significantly influence individuals' purchase intention by increasing their cultural intelligence. However, cultural distance was found to significantly mitigate the positive effects of both informational and socializing social media usage on cultural intelligence.

Findings

The research findings illustrated that the mechanism why and how social media usage exerts impacts on potential customers' purchase intention in social commerce, particularly in complicated cross-cultural environment. More specifically, social media usage, whether it is for informational or socializing purpose, exerts positive effects on users' purchase intention. The research findings partially keep in line with previous ones to some extent, which suggested that social media usage influences users' online psychological well-being and online intention

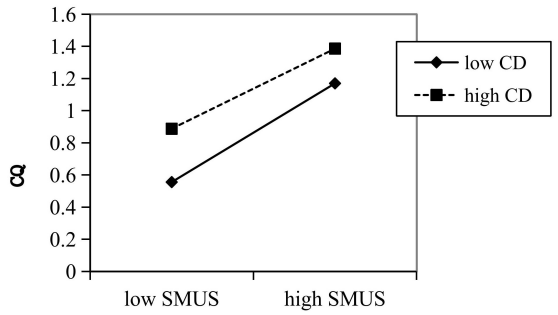


FIGURE 3 | Moderating effect of cultural distance on the relationship between informational social media usage and cultural intelligence.

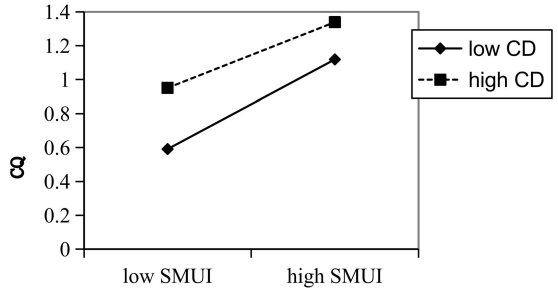


FIGURE 4 | Moderating effect of cultural distance on the relationship between socializing social media usage and cultural intelligence.

(Ellison et al., 2007; Hu et al., 2017, 2018; Priyo et al., 2019; Wang et al., 2021).

Further, the research unveiled how purchase intention in social commerce takes place in cross-cultural environment. The study has identified that when individuals are exposed to culturally diversified environment, social media usage helps them to develop their cultural intelligence, thereby enhancing users' purchase intention. Further, cultural intelligence was found to partially mediate the relationships between two dimensions of social media usage and users' purchase intention. The research findings keep compatible with prior research claiming that cultural intelligence plays as a conduit whereby social media usage functions on users' motivation, attitudes and behaviors, particularly in a culturally diversified business environment (Hu et al., 2017). In addition, the research discovered the role of cultural distance attenuating the effects exerted by both two dimensions of social media usage on cultural intelligence. The result indicates that cultural distance undermines the positive influences of social media usage on users' cultural intelligence. In particular, cultural distance exerts more negative effects on social media usage for informational purpose than for socializing purpose. This might be explained that socializing social media usage has been regarded as an efficacious way to facilitate establishment of interpersonal trust and mutual understandings, which may contribute to diluting the negative effects emanating from cultural distance (Lu et al., 2016). As such, the current research has confirmed that cultural distance impairs the functionalities of social media usage to develop individual's cultural intelligence in the environment with cultural novelty, thereby influencing users' purchase intention from their surfed social commerce websites. The research findings also match well with previous literature stating that cultural distance is a boundary element deterring an individual's psychological well-being and personal development in cross-cultural settings (Froese and Peltokorpi, 2011; Ng, 2013).

Research Implications

Based on social learning theory, an interesting research model was proposed for the current cross-cultural social commerce study. Accordingly, research hypotheses were tested with empirical evidence. The research findings make important contributions to the extant literature researching social commerce in several aspects.

First of all, the study partially unveils the underlying mechanism of social commerce which is supposed to be conducted in the environment transcending borders with cultural novelty. As for as we know, this effort is among the first trials investigating the functioning role of social media usage on users' purchase intention, particularly with a cross-cultural background for social commerce operation. Although tremendous efforts have been made to uncover the determinants influencing users' online intention for social commerce behaviors (Stewart, 2003; Liang and Turban, 2011; Chen and Shen, 2015), there is paucity of efforts endeavored to examine how two aspects of social media usage correlate with users' purchase intention for their social commerce real practices. The research findings confirm that two aspects of social media usage facilitate users' purchase intention on social commerce websites, which advances our understandings of social media and extends social commerce research to a much more complex cross-cultural environment. Undoubtedly, the research supplements the nomological network of the extant social commerce research.

Secondly, the current research also integrates cross-cultural studies into the social commerce research field by identifying that cultural intelligence plays different roles mediating the effects of two different dimensions of social media usage respectively on users' purchase intention in social commerce practices. Although researchers have identified the significant role of cultural intelligence enhancing expatriates' cross-cultural adaptation, personal development (Li and Tsai, 2015; Hu et al., 2020), few previous studies have identified how individual's cultural intelligence functions in social commerce operations. The research findings confirm the important role of cultural intelligence mediating the effects of social media usage on users' purchase intention as an effective conduit. It is worthy to note that cultural intelligence functions differently when mediating the relationships between two dimensions of social media usage and users purchase intention respectively. The research findings shed light on the different mechanisms whereby social media usage influences users' purchase intention in culturally diversified social commerce environment. The findings keep compatible with previous studies revealing the different mediating roles of cultural intelligence in a novel environment (Hu et al., 2017,

2021a). In addition, the current research is also among the first trials investigating how one contextual factor confounds another contextual factor to function on social commerce. To be more specific, cultural distance was found to act as a boundary condition to prevent social media usage from developing users' cultural intelligence in a culturally diversified social commerce environment. More specifically, social media usage demonstrates more positive associations with expatriates' cultural intelligence in a less culturally distant environment where the environment is less complex due to relatively enjoyable cultural similarity. It is also interesting to note that cultural distance exerts more negative effects on socializing social media usage than on informational social media usage, which contributes to the scholarly efforts of unveiling a complete picture of how cultural distance influences social media usage in a cross-cultural context. And those research findings contribute well to our understandings pertaining to how to enhance users' purchase intention on social commerce websites in a foreign environment with cultural novelty.

Thirdly, the research also extends the application of social learning theory to the social commerce research. Originated to serve psychological research, the social learning theory demonstrates its power to explicate individuals' skill and knowledge acquisition in specific contexts (Tu, 2000; Yi and Davis, 2003; Hu et al., 2020). However, little research has ever employed this theoretical lens to illuminate individuals' behaviors in social commerce research. The current research extended the theory to delineate users' cognitive and emotional changes in fostering their purchase intention in culturally diversified social commerce environment. The research findings keep consistent with the proposal that social learning underpins social commerce application as a fundamental social psychology (Chen et al., 2017).

The current investigation also provides three insights pertaining to practical and managerial practices directed at promoting social commerce operations, particularly international business operations.

First, much attention should be paid to the significant enabling roles of social media usage in promoting users' purchase intention on social commerce sites when international consumers are exposed to a culturally diversified business environment. Social media usage serves to positively influence users' purchase intention on social commerce websites (Pradana and Ha, 2021). Measures encouraging potential consumers to use social media to surf the company's social commerce websites or communicate with other consumers who are loyal to the company's social commerce should be considered. Well-designed company's social commerce websites with valuable and credible information posted by the company itself or shared by other consumers would facilitate potential consumers' understandings about products and marketing strategies originated from cultures. Further, convenient communicative channels designed on social commerce websites serve better interactivity between community members and contribute to stickness of the company's websites. Successful communications through the platform of social commerce websites contribute to the establishment of robust interpersonal relationships between consumers and offer avenues to spread products' word-of-mouth (Hariguna et al., 2020). And this will accumulate interests and motivations of potential consumers and increase their purchase intention on social commerce websites.

Second, given that cultural intelligence plays the role of mediating the relationships between social media usage and users' purchase intention on s-commerce sites, individuals with insufficient cultural intelligence tend to be less motivated to leverage social media to seek for necessary business information, fail to establish or maintain interpersonal relationships for intimate communications about business products, and finally lose their interests and purchase intention on social commerce websites. As such, any culture-related programs or activities including well-designed social commerce websites with necessary cultural knowledge to promote potential consumers' knowledge about foreign products' reflections of another country's or region's culture should be regarded as complementary and effective measures for appropriate marketing strategies. Those measures are supposed to develop cultural intelligence of targeted international consumers, thereby generating their purchase intention for online shopping behaviors. For international consumers, their capability to understand the culture nourishing the products and relevant business promotions seems to be crucial to determine their purchase decision. In this regard, measures promoting the cultural knowledge closely related to the products are strongly encouraged.

Finally, one important insight for practitioners to conduct successful international business operations in a foreign culture is that people with different cultural values should be addressed appropriately (Wankel, 2016). Thus cultural distance should be given full consideration during the process of product designing and promotion in a cross-cultural environment. Any marketing strategies promoting a company public image or culturally rooted products should be meticulously designed with serious consideration about the cultural dissimilarities between countries. Taken the event of famous Italian D and G Company as an example, which happened in November 2018, the company's advertising activity for a fashion show organized in Shanghai, China was regarded as an insulting conflict against Chinese tradition and culture. This event gained rapid fermentation through social media and the company was confronted with its most serious marketing crisis in China. Thus being fully aware of potential risks brought by cultural distance is the very first step for practitioners to deliver products' information on social media. Appropriate marketing approaches to convey and interpret culturally rooted products help to trigger a customer's interest, intrinsic motivation, and cultural identification to develop their cultural intelligence and render their purchase intention on s-commerce sites. In addition, marketing strategies designed to seek cultural similarities or similar inherent values between target and home countries may attenuate the negative effects brought by cultural differences, which contributes to the user's cultural intelligence and helps to generate more purchase intention.

In addition to the theoretical and practical implications, the current study makes relevant contributions to society as well. First of all, the study confirms the enabling roles of social media in

social commerce. This confirms the mutually melt and integrative relationships between information technology advancement and business prosperity. Development of information technology changes the operation modes and enhances the success of every business. In the era of digitalization, business and technology should be further combined and integrated. Secondly, in the process of world's economic globalization, businesses are supposed to confront increasing political, economic and cultural challenges. In response to those encounters, business development should consider global layout to avoid unexpected risks. Given the limited capacities of domestic markets, international approach could be one part of alternatives to increase business efficiency. The current study indicates that culture-relevant elements should be seriously considered for international business operations. Increasing cultural awareness and knowledge of international consumers is an effective conduit to enhance their purchase intention in cross-cultural social commerce. In this regard, cultural knowledge sharing between nations contributes to business success. Last but not the least, cultural distance, as indicated by previous research, has been identified as a dominating factor deterring interpersonal communications between individuals from different cultural backgrounds. Cultures demonstrate their effects from both positive and negative effects on human's lives. On the one hand, prosperity in cultures finally constitutes one part of human's glorious civilization. On the other hand, different cultures bring about obstacles to people's communications and even social stability. In this regard, measures mitigating negative effects of cultural differences should be taken to avoid potential cultural conflicts and help people with different cultural identifications to enjoy other nations' cultural glamor. And this finally vitalizes human's development in every domain including business operations.

LIMITATIONS AND FUTURE AVENUES

Regardless of the above-mentioned important contributions, the study also has limitations that could be addressed as prospective research efforts. First, the selected sample for the current study could be problematic. Albeit we selected at random appropriate international students samples studying at Chinese universities, the majority of students participating in the survey were from Africa and Asia. The group of international students can be regarded as important international social commerce consumers, whereas there might be differences in other groups of international consumers. Further, he sampling areas limitation may narrow the generalization and application of research findings in other parts of the world. Therefore, researchers should be cautious to extend the research findings to other groups of international consumers. More empirical research should be replicated to re-examine the findings with samples from other countries.

Second, the research results indicate that cultural intelligence plays the role of mediating the relationships between social media usage and users' purchase intention in culturally diversified social commerce environment. However, the current study only chose the key research construct in cross-cultural studies, cultural intelligence as the mediator aiming to unveil how social media usage exerts impact on consumers' purchase intention. In addition, cultural intelligence was found only to partially mediate the relationships. Given the more complexity and uncertainty existing in culturally diversified social commerce business environment compared to a singular cultural background, more empirical research deems to be supplementary to further investigate any other mediators and situational factors to explicate how social media usage influences users' purchase intention.

Finally, the research method to examine the proposed framework could be another problem. The adopted cross-sectional design only presents correlative, not causal relations between variables. For instance, the underlying mechanism of social media usage working on cultural intelligence could be reverse in such a way that individuals who possess sufficient cultural intelligence are supposed to rely on social media more heavily because they are more aware of the importance of social media in a foreign environment to help them develop personal capacities. Notwithstanding the conclusion reached with a longitudinal research design indicating that social media usage at first wave results in second-wave cultural intelligence (Hu et al., 2018), future empirical investigations are still necessary to offer a much clearer picture about antecedents determining users' social commerce behaviors or intention in the future.

CONCLUSION

Adopting social learning theory as the theoretical lens, the current research conducted an in-depth empirical study about cross-cultural social commerce. In particular, we adopted a multidisciplinary perspective by combining social commerce research and cross-cultural studies to investigate the underlying mechanism whereby social media usage exerts effects on users' purchase intention via cultural intelligence, while also considering the attenuating effects of cultural distance. The current research serves as one of the first trials to call for researching social commerce with an interdisciplinary approach in the future efforts.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

SH did the conceptualization, performed the methodology, and wrote the original draft. ZZ carried out the software, visualized and investigated the data, and wrote, reviewed, and edited the manuscript. Both authors contributed to the article and approved the submitted version.

FUNDING

This research was originally presented at the 19th Wuhan International Conference on E-business held in July 2020, China and financially supported by 2019' National Education Sciences Planning Project of China (BBA190019), Anhui Provincial Natural Science Foundation (1908085MG238), Anhui Provincial Teaching Reforming Project (2020jyxm0206), and Anhui University of Technology Research Project (2020syzm01).

REFERENCES

Adiandari, A. M. (2022). Financial performance innovation since digital technology entered Indonesian MSMEs. *Int. J. Appl. Inform. Manag.* 2, 50–58.

Algharabat, R. S., and Rana, N. P. (2021). Social commerce in emerging markets and its impact on online community engagement. *Inform. Syst. Front.* 23, 1499–1520. doi: 10.1007/s10796-020-10041-4

AlMazrouei, H., and Zacca, R. (2021). Cultural intelligence as a predictor of expatriate managers turnover intention and creative self-efficacy. *Int. J. Organ. Anal.* 29, 59–77. doi: 10.1108/ijoa-10-2019-1904

Alnoor, A., Al-Abrrow, H., Al Halbusi, H., Khaw, K. W., Chew, X., Al-Maatoq, M., et al. (2022). Uncovering the antecedents of trust in social commerce: an application of the non-linear artificial neural network approach. *Competitiveness Rev.* Epub ahead of print.

Ang, S., and Van Dyne, L. (2015). *Handbook of Cultural Intelligence*. London: Routledge.

Ang, S., Van Dyne, L., Koh, C., Ng, K. Y., Templer, K. J., Tay, C., et al. (2007). Cultural intelligence: its measurement and effects on cultural judgment and decision making, cultural adaptation, and task performance. *Manag. Organ. Rev.* 3, 335–371. doi: 10.1111/j.1740-8784.2007.00082.x

Aral, S., Dellarocas, C., and Godes, D. (2013). Social media and business transformation: a framework for research. *Inform. Syst. Res.* 24, 3–13. doi: 10.1287/isre.1120.0470

Bagozzi, R. P. (1981). Evaluating structural equation models with unobservable variables and measurement error: a comment. *J. Mark. Res.* 18, 375–381. doi: 10.2307/3150979

Baltacı, A. (2017). Relations between prejudice, cultural intelligence and level of entrepreneurship: a study of school principals. *Int. Electron. J. Elem. Educ.* 9, 645–666.

Bandura, A. (1978). Social learning theory of aggression. *J. Commun.* 28, 12–29. doi: 10.1111/j.1460-2466.1978.tb01621.x

Black, J. S., Mendenhall, M., and Oddou, G. (1991). Toward a comprehensive model of international adjustment: an integration of multiple theoretical perspectives. *Acad. Manag. Rev.* 16, 291–317. doi: 10.5465/amr.1991.4278938

Bonebrake, K. (2002). College students' internet use, relationship formation, and personality correlates. *Cyberpsychol. Behav.* 5, 551–557. doi: 10.1089/109493102321018196

Bugshan, H., and Attar, R. W. (2020). Social commerce information sharing and their impact on consumers. *Technol. Forecast. Soc. Chang.* 153:119875. doi: 10.1016/j.techfore.2019.119875

Busalim, A. H., and Hussin, A. R. C. (2016). Understanding social commerce: a systematic literature review and directions for further research. *Int. J. Inform. Manag.* 36, 1075–1088. doi: 10.1016/j.ijinfomgt.2016.06.005

Chavis, A. M. (2011). Social learning theory and behavioral therapy: considering human behaviors within the social and cultural context of individuals and families. *Soc. Work Public Health* 26, 471–481. doi: 10.1080/19371918.2011.591629

Chawla, N., and Kumar, B. (2021). E-commerce and consumer protection in india: the emerging trend. *J. Bus. Ethics* 1, 1–24. doi: 10.1007/s10551-021-04884-3

Chen, A., Lu, Y., and Wang, B. (2017). Customers' purchase decision-making process in social commerce: a social learning perspective. *Int. J. Inform. Manag.* 37, 627–638. doi: 10.1016/j.ijinfomgt.2017.05.001

Chen, G., Kirkman, B. L., Kim, K., Farh, C. I. C., and Tangirala, S. (2010). When does cross-cultural motivation enhance expatriate effectiveness? a multilevel investigation of the moderating roles of subsidiary support and cultural distance. *Acad. Manag. J.* 53, 1110–1130. doi: 10.5465/amj.2010.54533217

Chen, J., and Shen, X. L. (2015). Consumers' decisions in social commerce context: an empirical investigation. *Decis. Support Syst.* 79, 55–64. doi: 10.1016/j.dss.2015.07.012

Chen, N., Chao, M. C., Xie, H., and Tjosvold, D. (2018). Transforming cross-cultural conflict into collaboration: the integration of western and eastern values. *Cross Cult. Strateg. Manag.* 25, 70–95. doi: 10.1108/ccsm-10-2016-0187

Chen, X. P., Liu, D., and Portnoy, R. (2012). A multilevel investigation of motivational cultural intelligence, organizational diversity climate, and cultural sales: evidence from US real estate firms. *J. Appl. Psychol.* 97, 93–106. doi: 10.1037/a0024697

Chen, X. P., and Peng, S. (2008). Guanxi dynamics: shifts in the closeness of ties between Chinese coworkers. *Manag. Organ. Rev.* 4, 63–80. doi: 10.1111/j.1740-8784.2007.00078.x

Curty, R. G., and Zhang, P. (2013). Website features that gave rise to social commerce: a historical analysis. *Electron. Commer. Res. Appl.* 12, 260–279. doi: 10.1016/j.elerap.2013.04.001

Deng, Z., and Wang, Z. (2016). Early-mover advantages at cross-border business-to-business e-commerce portals. *J. Bus. Res.* 69, 6002–6011. doi: 10.1016/j.jbusres.2016.05.015

Earley, P. C., and Ang, S. (2003). *Cultural Intelligence: Individual Interactions across Cultures*. Stanford, CA: Stanford University Press.

Ellison, N. B., Steinfield, C., and Lampe, C. (2007). The benefits of Facebook "friends:" social capital and college students' use of online social network sites. *J. Comput. Med. Commun.* 12, 1143–1168. doi: 10.1111/j.1083-6101.2007.00367.x

English, A. S., Zhang, Y. B., and Tong, R. (2021). Social support and cultural distance: sojourners' experience in China. *Int. J. Intercult. Relat.* 80, 349–358. doi: 10.1016/j.ijintrel.2020.10.006

Fornell, C., and Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *J. Mark. Res.* 18, 39–50. doi: 10.2307/3151312

Friedman, T. L. (2006). *The World is Flat: A Brief History of the Twenty-First Century (Rev. ed.)*. New York, NY: Farrar, Straus and Giroux.

Friedrich, T., Schlauderer, S., and Overhage, S. (2019). The impact of social commerce feature richness on website stickiness through cognitive and affective factors: an experimental study. *Electron. Commer. Res. Appl.* 36:100861. doi: 10.1016/j.elerap.2019.100861

Frijns, B., and Garel, A. (2021). The effect of cultural distance between an analyst and a CEO on analysts' earnings forecast performance. *Econ. Lett.* 205:109957. doi: 10.1016/j.econlet.2021.109957

Froese, F. J., and Peltokorpi, V. (2011). Cultural distance and expatriate job satisfaction. *Int. J. Intercult. Relat.* 35, 49–60. doi: 10.1016/j.ijintrel.2010.10.002

Fu, J.-R., Lu, I. W., Chen, J. H. F., and Farn, C.-K. (2020). Investigating consumers' online social shopping intention: an information processing perspective. *Int. J. Inform. Manag.* 54:102189. doi: 10.1016/j.ijinfomgt.2020.102189

Galchenko, I., and van de Vijver, F. J. R. (2007). The role of perceived cultural distance in the acculturation of exchange students in Russia. *Int. J. Intercult. Relat.* 31, 181–197. doi: 10.1016/j.ijintrel.2006.03.004

Gassmann, O., Zeschky, M., Wolff, T., and Stahl, M. (2010). Crossing the industry-line: breakthrough innovation through cross-industry alliances with 'non-suppliers'. *Long Range Plann.* 43, 639–654. doi: 10.1016/j.lrp.2010.06.003

Ghahtarani, A., Sheikhmohammady, M., and Rostami, M. (2020). The impact of social capital and social interaction on customers' purchase intention, considering knowledge sharing in social commerce context. *J. Innov. Knowl.* 5, 190–198.

Glass, C. R., and Westmont, C. M. (2014). Comparative effects of belongingness on the academic success and cross-cultural interactions of domestic and international students. *Int. J. Intercult. Relat.* 39, 152–163.

Guan, L., Liang, H., and Zhu, J. J. H. (2022). Predicting reposting latency of news content in social media: a focus on issue attention, temporal usage pattern, and information redundancy. *Comput. Hum. Behav.* 127:107080. doi: 10.1016/j.chb.2021.107080

Guang, X., and Charoensukmongkol, P. (2022). The effects of cultural intelligence on leadership performance among Chinese expatriates working in Thailand. *Asian Bus. Manag.* 21, 106–128. doi: 10.1057/s41291-020-00112-4

Guillaume, Y. R. F., Knippenberg, D. V., and Brodbeck, F. C. (2014). Nothing succeeds like moderation: a social self-regulation perspective on cultural dissimilarity and performance. *Acad. Manag. J.* 57, 1284–1308. doi: 10.5465/amj.2013.0046

Gutama, D. H., Umami, I., and Saputro, P. H. (2021). Analysis of the effect of website sales quality on purchasing decisions on e-commerce websites. *Int. J. Inform. Inform. Syst.* 4, 71–81. doi: 10.47738/ijiis.v4i1.79

Hair, J. F. Jr., Black, W. C., Babin, B. J., and Anderson, R. E. (2020). *Multivariate Data Analysis*, 8th Edn. Upper Saddle River, NJ: Prentice Hall.

Hajli, N. (2015). Social commerce constructs and consumer's intention to buy. *Int. J. Inform. Manag.* 35, 183–191. doi: 10.1016/j.ijinfomgt.2014.12.005

Hariguna, T., and Rachmawati, V. (2019). Community opinion sentiment analysis on social media using naive bayes algorithm methods. *Int. J. Inform. Inform. Syst.* 2, 33–38. doi: 10.47738/ijiis.v2i1.11

Hariguna, T., Sukmana, H. T., and Kim, J. I. (2020). Survey opinion using sentiment analysis. *J. Appl. Data Sci.* 1, 35–40. doi: 10.47738/jads.v1i1.10

Hoppner, J. J., Griffith, D. A., and White, R. C. (2015). Reciprocity in relationship marketing: a cross-cultural examination of the effects of equivalence and immediacy on relationship quality and satisfaction with performance. *J. Int. Mark.* 23, 64–83. doi: 10.1509/jim.15.0018

Hossain, M. S., Rahman, M. F., and Zhou, X. (2021). Impact of customers' interpersonal interactions in social commerce on customer relationship management performance. *J. Contemp. Mark. Sci.* 4, 161–181. doi: 10.1108/jcmars-12-2020-0050

Hu, S., Gu, J., Liu, H., and Huang, Q. (2017). The moderating role of social media usage in the relationship among multicultural experiences, cultural intelligence, and individual creativity. *Inform. Technol. People* 30, 265–281. doi: 10.1108/itp-04-2016-0099

Hu, S., Hu, L., and Wang, G. (2021a). Moderating role of addiction to social media usage in managing cultural intelligence and cultural identity change. *Inform. Technol. People* 34, 704–730. doi: 10.1108/itp-10-2019-0518

Hu, S., Hu, L., Wu, J., and Wang, G. (2021b). Social media usage and international expatriate's creativity: an empirical research in cross-cultural context. *Hum. Manag. Syst.* 40, 197–209. doi: 10.3233/hsm-200965

Hu, S., Liu, H., and Gu, J. (2018). What role does self-efficacy play in developing cultural intelligence from social media usage? *Electron. Commer. Res. Appl.* 28, 172–180. doi: 10.1016/j.elerap.2018.01.009

Hu, S., Liu, H., Zhang, S., and Wang, G. (2020). Proactive personality and cross-cultural adjustment: roles of social media usage and cultural intelligence. *Int. J. Intercult. Relat.* 74, 42–57. doi: 10.1016/j.ijintrel.2019.10.002

Huang, Z., and Benyoucef, M. (2013). From e-commerce to social commerce: a close look at design features. *Electron. Commer. Res. Appl.* 12, 246–259. doi: 10.1016/j.elerap.2012.12.003

Huang, Z., and Benyoucef, M. (2015). User preferences of social features on social commerce websites: an empirical study. *Technol. Forecast. Soc. Chang.* 95, 57–72. doi: 10.1016/j.techfore.2014.03.005

Hughes, D. J., Rowe, M., Batey, M., and Andrew, L. (2012). A tale of two sites: Twitter vs. Facebook and the personality predictors of social media usage. *Comput. Hum. Behav.* 28, 561–569. doi: 10.1016/j.chb.2011.11.001

Hur, K., Kim, T. T., Karatepe, O. M., and Lee, G. (2017). An exploration of the factors influencing social media continuance usage and information sharing intentions among Korean travelers. *Tour. Manag.* 63, 170–178. doi: 10.1016/j.tourman.2017.06.013

Hwang, I. J., Lee, B. G., and Kim, K. Y. (2014). Information asymmetry, social networking site word of mouth, and mobility effects on social commerce in Korea. *Cyberpsychol. Behav. Soc. Netw.* 17, 117–124. doi: 10.1089/cyber.2012.0566

Illeris, K. (2003). Towards a contemporary and comprehensive theory of learning. *Int. J. Lifelong Educa.* 22, 396–406. doi: 10.1080/02601370304837

Imai, L., and Gelfand, M. J. (2010). The culturally intelligent negotiator: the impact of cultural intelligence (CQ) on negotiation sequences and outcomes. *Organ. Behav. Hum. Decis. Process.* 112, 83–98. doi: 10.1016/j.obhdp.2010.02.001

Jami Pour, M., Hosseinzadeh, M., and Mansouri, N. S. (2022). Challenges of customer experience management in social commerce: an application of social network analysis. *Internet Res.* 32, 241–272. doi: 10.1108/intr-01-2021-0076

Jen, L., and Lin, Y. (2021). A brief overview of the accuracy of classification algorithms for data prediction in machine learning applications. *J. Appl. Data Sci.* 2, 84–92. doi: 10.47738/jads.v2i3.38

Jin, X., Chen, X., and Zhou, Z. (2022). The impact of cover image authenticity and aesthetics on users' product-knowing and content-reading willingness in social shopping community. *Int. J. Inform. Manag.* 62:102428. doi: 10.1016/j.ijinfomgt.2021.102428

Karatas, K., and Arpaci, I. (2021). The mediating role of tolerance in the relationship between cultural intelligence and xenophobia. *Asia Pac. Educ. Rev.* 22, 119–127. doi: 10.1007/s12564-021-09675-z

Kauffman, R. J., Kim, K., Lee, S. Y., Hoang, A. P., and Ren, J. (2017). Combining machine based and econometrics methods for policy analytics insights. *Electron. Commer. Res. Appl.* 25, 115–140. doi: 10.1016/j.elerap.2017.04.004

Khaola, P. P., Musiiwa, D., and Rambe, P. (2022). The influence of social media usage and student citizenship behaviour on academic performance. *Int. J. Manag. Educ.* 20:100625. doi: 10.1002/ijop.12200

Kim, S., and Park, H. (2013). Effects of various characteristics of social commerce (s-commerce) on consumers' trust and trust performance. *Int. J. Inform. Manag.* 33, 318–332. doi: 10.1016/j.ijinfomgt.2012.11.006

Kim, Y. J., and Dyne, L. V. (2012). Cultural intelligence and international leadership potential: the importance of contact for members of the majority. *Appl. Psychol. Int. Rev.* 61, 272–294. doi: 10.1111/j.1464-0597.2011.00468.x

Ko, H. C. (2018). Social desire or commercial desire? The factors driving social sharing and shopping intentions on social commerce platforms. *Electron. Commer. Res. Appl.* 28, 1–15. doi: 10.1016/j.elerap.2017.12.011

Kotabe, M., Srinivasan, S. S., and Aulakh, P. S. (2002). Multinationality and firm performance: the moderating role of RandD and marketing capabilities. *J. Int. Bus. Stud.* 33, 79–97. doi: 10.1057/palgrave.jibs.8491006

Lankton, N. K., McKnight, D. H., Wright, R. T., and Thatcher, J. B. (2016). Research note—using expectation disconfirmation theory and polynomial modeling to understand trust in technology. *Inform. Syst. Res.* 27, 197–213. doi: 10.1287/isre.2015.0611

LaRose, R. (2009). "Social cognitive theories of media selection," in *Media Choice: A Theoretical and Empirical Overview*, ed. T. Hartmann (New York, NY: Routledge), 10–31.

Lee, L. Y., and Sukoco, B. M. (2010). The effffects of cultural intelligence on expatriate performance: the moderating effffects of international experience. *Int. J. Hum. Resour. Manag.* 21, 963–981. doi: 10.1080/09585191003783397

Leong, L.-Y., Hew, T.-S., Ooi, K.-B., and Chong, A. Y.-L. (2020). Predicting the antecedents of trust in social commerce – a hybrid structural equation modeling with neural network approach. *J. Bus. Res.* 110, 24–40. doi: 10.1016/j.jbusres.2019.11.056

Li, C., and Tsai, W. H. S. (2015). Social media usage and acculturation: a test with Hispanics in the U.S. *Comput. Hum. Behav.* 45, 204–212. doi: 10.1016/j.chb.2014.12.018

Liang, T. P., and Turban, E. (2011). Introduction to the special issue on social commerce: a research framework for social commerce. *Int. J. Electron. Commer.* 16, 5–13. doi: 10.2753/jec1086-4415160201

Lickel, B., Hamilton, D. L., Wieczorkowska, G., Lewis, A., Sherman, S. J., and Uhles, A. N. (2000). Varieties of groups and the perception of group entitativity. *J. Pers. Soc. Psychol.* 78, 223–246. doi: 10.1037//0022-3514.78.2.223

Lin, C. S., Kuo, F. Y., and Hung, C. Y. (2021). Exploring social media use and civic engagement on the discussion of antinuclear issue. *Int. J. Appl. Inform. Manag.* 2, 90–96. doi: 10.47738/ijaim.v2i2.31

Lin, X., Li, Y., and Wang, X. (2017). Social commerce research: definition, research themes and the trends. *Int. J. Inform. Manag.* 37, 190–201. doi: 10.1016/j.ijinfomgt.2016.06.006

Liu, A., Lu, C., and Wang, Z. (2021). Does cultural distance hinder exports?: a comparative study of China and the United States. *Econ. Model.* 105, 105668. doi: 10.1016/j.econmod.2021.105668

Lorenzo, O., Kawalek, P., and Ramdani, B. (2012). Enterprise applications diffusion within organizations: a social learning perspective. *Inform. Manag.* 49, 47–57. doi: 10.1016/j.im.2011.10.005

Lu, B., Fan, W., and Zhou, M. (2016). Social presence, trust, and social commerce purchase intention: an empirical research. *Comput. Hum. Behav.* 56, 225–237. doi: 10.1016/j.chb.2015.11.057

Ma, Z. (2010). The SINS in business negotiations: explore the cross-cultural differences in business ethics between Canada and China. *J. Bus. Ethics* 91, 123–135. doi: 10.1007/s10551-010-0571-5

MacNab, B. R., and Worthley, R. (2012). Individual characteristics as predictors of cultural intelligence development: the relevance of self-effiffifficacy. *Int. J. Intercult. Relat.* 36, 62–71. doi: 10.1016/j.ijintrel.2010.12.001

Matsuda, T. (2022). Research on new regional creation business model utilizing social network and crowdfunding. *Int. J. Appl. Inform. Manag.* 2, 1–12. doi: 10.47738/ijaim.v2i1.22

McCort, D. J., and Malhotra, N. K. (1993). Culture and consumer behavior: toward an understanding of cross-cultural consumer behavior in international marketing. *J. Int. Consum. Mark.* 6, 91–127. doi: 10.1300/j046v06n02_07

Mendenhall, M., and Oddou, G. (1985). The dimensions of expatriate acculturation: a review. *Acad. Manag. Rev.* 10, 39–47. doi: 10.5465/amr.1985.4277340

Mikalef, P., Giannakos, M. N., and Pappas, I. O. (2017). Designing social commerce platforms based on consumers' intentions. *Behav. Inform. Technol.* 36, 1308–1327. doi: 10.1080/0144929x.2017.1386713

Ministry of Education of People's Republic of China (2018). *Statistics of International Students Studying in China 2018*. Available online at: http://www.moe.gov.cn/jyb_xwfb/gzdt_gzdt/s5987/201904/t20190412_377692.html (accessed April 12, 2019).

Mohammed, E., and Qhal, A. (2020). Knowledge management strategy by means of virtualization in covid-19. *J. Appl. Data Sci.* 1, 41–53. doi: 10.7759/cureus.15699

Nacar, R., and Ozdemir, K. (2022). "From commerce to E-commerce and social commerce: how global? How local?," in *Industry 4.0 and Global Businesses*, ed. E. Yakut (Bingley: Emerald Publishing Limited), 95–109. doi: 10.1108/978-1-80117-326-120211007

Ng, C. S. P. (2013). Intention to purchase on social commerce websites across cultures: a cross-regional study. *Inform. Manag.* 50, 609–620. doi: 10.1016/j.im.2013.08.002

Osatuyi, B., Qin, H., Osatuyi, T., and Turel, O. (2020). When it comes to satisfaction … It depends: an empirical examination of social commerce users. *Comput. Hum. Behav.* 111:106413. doi: 10.1016/j.chb.2020.106413

Paireekreng, W., Osathanukroh, J., and Supasak, C. (2019). A study of influence factors for advertising on messaging applications towards mobile buyer's decision making. *Int. J. Inform. Inform. Syst.* 2, 82–90.

Pidduck, R. J., Shaffer, M. A., Zhang, Y., Cheung, S. S. Y., and Yunlu, D. G. (2022). Cultural intelligence: an identity lens on the influence of cross-cultural experience. *J. Int. Manag.* 28:100928. doi: 10.1016/j.intman.2022.100928

Podsakoff, P. M., Mackenzie, S. B., and Podsakoff, N. P. (2012). Sources of method bias in social science research and recommendations on how to control it. *Annu. Rev. Psychol.* 63, 539–569. doi: 10.1146/annurev-psych-120710-100452

Prabowo, N. A., Pujiarto, B., Wijaya, F. S., Gita, L., and Alfandy, D. (2021). Social network analysis for user interaction analysis on social media regarding e-commerce business. *Int. J. Inform. Inform. Syst.* 4, 95–102. doi: 10.47738/ijiis.v4i2.106

Pradana, M. G., and Ha, H. T. (2021). Maximizing strategy improvement in mall customer segmentation using K-means clustering. *J. Appl. Data Sci.* 2, 19–25.

Presbitero, A. (2018). Extraversion, openness to experience, and global career intention: the mediating role of cultural intelligence. *J. Employ. Couns.* 55, 104–114. doi: 10.1002/joec.12090

Priyo, J. S., Mohamad, B., and Adetunji, R. R. (2019). An examination of the effects of service quality and customer satisfaction on customer loyalty in the hotel industry. *Int. J. Supply Chain Manag.* 8, 653–663.

Rakhmansyah, M., Wahyuningsih, T., Srenggini, A. D., and Gunawan, I. K. (2022). Small and medium enterprises (SMEs) with SWOT analysis method. *Int. J. Appl. Inform. Manag.* 2, 47–54. doi: 10.47738/ijaim.v2i3.37

Schumann, J. H., Wangenheim, F. V., Stringfellow, A., Yang, Z., Blazevic, V., Praxmarer, S., et al. (2010). Cross-cultural differences in the effect of received word-of-mouth referral in relational service exchange. *J. Int. Mark.* 18, 62–80. doi: 10.1509/jimk.18.3.62

Shenkar, O. (2001). Cultural distance revisited: towards a more rigorous conceptualization and measurement of cultural differences. *J. Int. Bus. Stud.* 32, 1–17. doi: 10.1057/jibs.2011.40

Shin, D. H. (2013). User experience in social commerce: in friends we trust. *Behav. Inform. Technol.* 32, 52–67. doi: 10.1080/0144929x.2012.692167

Silbiger, A., Barnes, B. R., Berger, R., and Renwick, D. W. S. (2021). The role of regulatory focus and its influence on the cultural distance – adjustment relationship for expatriate managers. *J. Bus. Res.* 122, 398–410. doi: 10.1016/j.jbusres.2020.09.021

Soleimani, M. (2021). Buyers' trust and mistrust in e-commerce platforms: a synthesizing literature review. *Inform. Syst. Ebus. Manag.* doi: 10.1007/s10257-021-00545-0 [Epub ahead of print].

Stewart, K. J. (2003). Trust transfer on the world wide web. *Organ. Sci.* 14, 5–17. doi: 10.1287/orsc.14.1.5.12810

Stewart, S. M., and Bing, M. N. (2009). In the eyes of the beholder: a non-self-report measure of workplace deviance. *J. Appl. Psychol.* 94, 207–215. doi: 10.1037/a0012605

Stoermer, S., Davies, S., and Froese, F. J. (2021). The influence of expatriate cultural intelligence on organizational embeddedness and knowledge sharing: the moderating effects of host country context. *J. Int. Bus. Stud.* 52, 432–453. doi: 10.1057/s41267-020-00349-3

Sun, Y., Shao, X., Li, X., Guo, Y., and Nie, K. (2019). How live streaming influences purchase intentions in social commerce: an IT affordance perspective. *Electron. Commer. Res. Appl.* 37:100886. doi: 10.1016/j.elerap.2019.100886

Tang, L. (2017). Mine your customers or mine your business: the moderating role of culture in online word-of-mouth reviews. *J. Int. Mark.* 25, 88–110. doi: 10.1509/jim.16.0030

Tett, R. P., and Burnett, D. D. (2003). A personality trait-based interactionist model of job performance. *J. Appl. Psychol.* 88, 500–517. doi: 10.1037/0021-9010.88.3.500

Triandis, H. C. (2006). Cultural intelligence in organizations. *Group Organ. Manag.* 31, 20–26. doi: 10.1177/1059601105275253

Tu, C. H. (2000). On-line learning migration: from social learning theory to social presence theory in a CMC environment. *J. Netw. Comput. Appl.* 23, 27–37. doi: 10.1006/jnca.1999.0099

Van Heel, B., Lukic, V., and Leeuwis, E. (2011). *Cross-Border E-Commerce Makes the World Flatter*. Boston, MA: Boston Consulting Group.

Wang, C., Yang, Y. Y., and Chiang, M. H. (2021). Understanding users' attitude to social endorsement advertising of embarrassing product. *Int. J. Appl. Inform. Manag.* 1, 6–22.

Wang, C., Yang, Y. Y., and Lin, M. (2019). The influence of the privacy concern and social advertising type on the attitude and behavior. *Int. J. Inform. Inform. Syst.* 2, 75–81. doi: 10.2196/15455

Wang, E. S. T. (2010). Internet usage purposes and gender differences in the effects of perceived utilitarian and hedonic value. *Cyberpsychol. Behav. Soc. Netw.* 13, 179–183. doi: 10.1089/cyber.2009.0200

Wang, X., Lin, X., and Spencer, M. K. (2019). Exploring the effects of extrinsic motivation on consumer behaviors in social commerce: revealing consumers' perceptions of social commerce benefits. *Int. J. Inform. Manag.* 45, 163–175. doi: 10.1016/j.ijinfomgt.2018.11.010

Wang, Y., and Herrando, C. (2019). Does privacy assurance on social commerce sites matter to millennials? *Int. J. Inform. Manag.* 44, 164–177. doi: 10.1016/j.ijinfomgt.2018.10.016

Wankel, C. (2016). Developing cross-cultural managerial skills through social media. *J. Organ. Chang. Manag.* 29, 116–124. doi: 10.1108/jocm-11-2015-0225

Ward, C., Wilson, J., and Ronald, F. (2011). Assessing the predictive validity of cultural intelligence over time. *Pers. Individ. Differ.* 51, 138–142. doi: 10.1016/j.paid.2011.03.032

Weller, K. (2016). Trying to understand social media users and usage: the forgotten features of social media platforms. *Online Inform. Rev.* 40, 256–264. doi: 10.1108/oir-09-2015-0299

Wigand, R. T., Benjamin, R. I., and Birkland, J. L. H. (2008). "Web 2.0 and beyond: implications for electronic commerce," in *Proceedings of the 10th International Conference on Electronic Commerce, Innsbruck, Austria* (New York, NY: ACM Press).

Yadav, M. S., de Valck, K., Hennig-Thurau, T., Hoffman, D. L., and Spann, M. (2013). Social commerce: a contingency framework for assessing marketing potential. *J. Interact. Mark.* 27, 311–323. doi: 10.1016/j.intmar.2013.09.001

Yazdanshenas, M. (2021). Core self-evaluations and project managers' competencies: the moderating role of cultural intelligence. *J. Manag. Dev.* 40, 542–573. doi: 10.1108/jmd-01-2021-0031

Yeo, G. B., and Neal, A. (2004). A multilevel analysis of effort, practice and performance: effects of ability, conscientiousness and goal orientation. *J. Appl. Psychol.* 89, 231–247. doi: 10.1037/0021-9010.89.2.231

Yi, M. Y., and Davis, F. D. (2003). Developing and validating an observational learning model of computer software training and skill acquisition. *Inform. Syst. Res.* 14, 146–169. doi: 10.1287/isre.14.2.146.16016

Zhang, J., and Goodson, P. (2011). Acculturation and psychosocial adjustment of Chinese international students: examining mediation and moderation effects. *Int. J. Intercult. Relat.* 35, 614–627. doi: 10.1016/j.ijintrel.2010.11.004

Zhang, K. Z. K., and Benyoucef, M. (2016). Consumer behavior in social commerce: a literature review. *Decis. Support Syst.* 86, 95–108.

Zheng, X., Men, J., Xiang, L., and Yang, F. (2020). Role of technology attraction and parasocial interaction in social shopping websites. *Int. J. Inform. Manag.* 51:102043. doi: 10.1016/j.ijinfomgt.2019.102043

Zhou, L., Zhang, P., and Zimmermann, H. D. (2013). Social commerce research: an integrated view. *Electron. Commer. Res. Appl.* 12, 61–68. doi: 10.1016/j.elerap.2013.02.003

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Hu and Zhu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers | Frontiers in Psychology

ORIGINAL RESEARCH
published: 12 May 2022
doi: 10.3389/fpsyg.2022.881725

The Effects of Leader Emotional Intelligence, Leadership Styles, Organizational Commitment, and Trust on Job Performance in the Real Estate Brokerage Industry

Chun-chang Lee[1], Yei-shian Li[1], Wen-chih Yeh[2] and Zheng Yu[3]*

[1] Department of Real Estate Management, National Pingtung University, Pingtung, Taiwan, [2] Department of Real Estate Management, HungKuo Delin University of Technology, New Taipei City, Taiwan, [3] Department of Land Economics, National Chengchi University, Taipei, Taiwan

OPEN ACCESS

Edited by:
Francesca D'Errico,
University of Bari Aldo Moro, Italy

Reviewed by:
Bikrant Kesari,
Maulana Azad National Institute of Technology, India
Ernestina Lamponi,
Università Telematica Internazionale Uninettuno, Italy
Maria Luisa Giancaspro,
University of Bari Aldo Moro, Italy

****Correspondence:***
Chun-chang Lee
lcc@mail.nptu.edu.tw

Specialty section:
This article was submitted to Performance Science, a section of the journal Frontiers in Psychology

Received: *23 February 2022*
Accepted: *07 April 2022*
Published: *12 May 2022*

Citation:
Lee C-C, Li Y-S, Yeh W-C and Yu Z (2022) The Effects of Leader Emotional Intelligence, Leadership Styles, Organizational Commitment, and Trust on Job Performance in the Real Estate Brokerage Industry. Front. Psychol. 13:881725. doi: 10.3389/fpsyg.2022.881725

This study examines the effects of leader emotional intelligence, leadership styles (transformational and transactional), organizational commitment, and trust on job performance. A questionnaire was administered to the participants, who were real estate brokers in Kaohsiung City. Of the 980 questionnaires administered, 348 valid responses were received, indicating an effective response rate of 35.5%. Structural equation modeling was used for the analysis. The results show that leader emotional intelligence has a significant and positive effect on trust in supervisors, and transformational leadership and trust within a team have significant and positive effects on job performance. In addition, organizational commitment has a significant and positive effect on job performance. Trust within a team mediates the significant and positive effect of leader emotional intelligence on job performance. Although transactional leadership has no direct, positive, and significant effects on employee job performance, trust in the workplace shaped by a leader's leadership style will empower a team's spirit and boost their morale, thereby indirectly affecting their job performance in a positive manner.

Keywords: real estate brokers, leadership emotional intelligence, leadership style, organizational commitment, trust, job performance

INTRODUCTION

In today's increasingly complex, ever-changing and fiercely competitive business environments, leaders must constantly enhance the competitiveness of their organization as well as their employees' performance to survive and succeed. According to Masa'deh et al. (2016), employee job performance significantly and positively affects a company's sales, because the sum of the tasks completed by every employee is reflected in the revenue of the company. Employees who are more capable of achieving set organizational goals contribute more to the company's revenue. Therefore, it is crucial for organizations to implement measures that improve their employees' job performance.

To sustain a company's competitiveness, not only must supervisors motivate their employees to leverage their skills to achieve the company's goals and interests, but it is equally important for a company to foster employees' resonance with and loyalty toward organizational values.

Organizational assets, such as important techniques or information, are obtained by employees as part of their job. Employees with a high level of organizational commitment identify themselves as a proud member of their organization who enjoys their job. Therefore, they are more efficient at work, have lower turnover intentions, are less likely to leak or steal confidential company information due to their loyalty, and avoid actions that damage their company's reputation. In addition, employee–team-leader relationships are an indispensable component in business operations, especially regarding mutual trust. Team members who trust one another enhance their job performance through healthy interactions and teamwork (Jong and Elfring, 2010). Employees who comprehend the demands and work goals required by their supervisor and company trust their leader to reward them for their job performance. Employees who give timely and adequate assistance and support to their company in times of need generate trust toward their supervisor and company, which motivates them to work harder and achieve better performances (Schaubroeck et al., 2013).

Concerning the factors that affect job performance, in addition to trust and organizational commitment, it is also important to consider leader emotional intelligence, which affects trust and employee job performance. Emotional intelligence refers to an individual's array of abilities to identify, express, understand, and evaluate their own emotions as well as those of others (Schlechter and Strauss, 2008; Harms and Credé, 2010). Kotsou et al. (2019) described emotional intelligence as the ability to identify, express, understand, manage, and use emotions. Emotional intelligence significantly affects our health, relationships, and work and learning experiences. Wen et al. (2019) proposed two ways to explain emotional intelligence. The first describes emotional intelligence as a social capability in which one is able to monitor their own as well as others' emotions, develop emotional cognition-based information, and use this information to guide their thinking and actions. The second describes emotional intelligence as the capability to perceive one's own and others' emotions and initiate one's emotional management and interpersonal relations. According to Wong and Law (2002), interpersonal interactions are essential for team operation; team members who are unable to regulate their emotions are more likely to find themselves at odds with others. Emotionally intelligent leaders who are aware of their subordinates' emotional state help team members control their emotions and prevent conflicts between team members caused by overreactions (Chang et al., 2012; Liu et al., 2012). D'Errico et al. (2022) pointed out that leaders who possess humility and moral suasion facilitate a team to sustain their cooperative spirit and foster stronger within-team trust. Therefore, leader emotional intelligence does more than only allow employees to achieve better job performances, from an external viewpoint; it also stresses a leader's humility and moral suasion, as well as controls the leader's and their employees' emotions, from an internal viewpoint. When both the internal and external criteria are achieved, successful teamwork and better job performance can be attained easily. Furthermore, emotionally intelligent leaders take the initiative to establish a positive and harmonious team morale and strengthen trust among team members, thereby enhancing the job performance and organizational effectiveness of managers and subordinates alike (Shahhosseini et al., 2012).

In addition, leadership styles are a factor that influences an employee's trust in their supervisor as well as their job performance. Excellent leaders stimulate their employees to express their potential, actively enhance their job performance, and attend to their needs. On a psychological level, employees who resonate well with their leader's leadership develop attitudes characterized by trust; on a behavioral level, they adopt behaviors that benefit their company and proactively dedicate themselves to achieving organizational goals. Therefore, organizational effectiveness is reflected through a leader, and leaders need to modify their leadership styles according to external environmental and temporal changes. Different leadership styles have different effects on organizational performance as well as the job performance of individual employees. Transformational leadership and transactional leadership are the two most common leadership styles. Transformational leadership stresses how leaders influence changes in their employees' ideas and behaviors such that employees are committed to meeting organizational visions and goals. Yue et al. (2019) suggested that transformational leaders participate in goal setting and directly offer strategic directions for their organization while considering and meeting employee demands. They encourage employees to think from different perspectives and seize opportunities for personal growth. Transformational leaders provide instruction and training to help their employees adapt to new work environments seamlessly. These leaders motivate employees to fully engage and dedicate themselves to their job in order to improve their job performance (MacKenzie et al., 2001), thereby increasing their job satisfaction. The individualized support and motivation provided by transformational leaders make their employees feel like they are being cared for and not isolated or helpless. Subsequently, employees trust their leader' leadership and are more willing to put more effort into their job and attain a better job performance (Braun et al., 2013). In transactional leadership, the leader–employee relationship is established on the grounds of transactions. Leaders set well-defined work goals and characteristics and implement rigorous reward or punishment schemes so that employees comprehend the benefits of carrying out their work tasks in accordance with their leader's demands. Employees receive the rewards they desire by accomplishing their superior's expected objectives. Transactional leaders need to set clear work standards for employees so they can understand the scope of their work and the basis for being rewarded or punished (Bass et al., 2003). These standards, in turn, generate employees' trust in their leaders, causing employees to believe that high performers are rewarded whereas low performers or those who make errors are corrected or punished (Asencio and Mujkic, 2016). Jung and Avolio (2000) indicated that trust in supervisors mediates the indirect and positive effect of transformational leadership and transactional leadership on job performance. In the real estate industry, real estate brokers are classified as high-ranking brokers (high commission but no base salary) and typical brokers (low commission but high base salary). Therefore, branch managers who wish to understand the performance and development of each employee should leverage

their own unique leader emotional intelligence (management of one's and others' emotions) to effectively formulate internal management mechanisms. To this end, transactional leadership should be applied when managing high-ranking brokers. Leaders should verify and define their subordinates' roles, require them to achieve set goals, understand how the subordinates perform their roles to achieve specified results, and use performance as an ultimate indicator of achievement. Typical brokers place more emphasis on the individual differences between employees, and leaders affect the values and motivations of employees. In addition to motivating their subordinates to attain transactional objectives, the goal of leaders is to arouse their subordinates' self-actualization, as opposed to immediately meeting self-interests.

Previous studies on job performance have mostly focused on employee satisfaction, organizational structure, and organizational commitment. For instance, Lee and Shen (2008) explored the effects of organizational structure, employee satisfaction, and organizational commitment on the job performance of real estate brokers. Hou (2012) examined the effects of organizational impartiality, organizational commitment, and trust on sugar company employees' willingness to share knowledge as well as their job performance. Chang (2016) examined how organizational climate, organizational trust, leadership styles, and internal control affects employees' organizational commitment and job performance. MacKenzie et al. (2001) discussed how job performance is affected by transformational leadership, transactional leadership, trust in supervisors, and role ambiguity. Schlechter and Strauss (2008) studied the relationships between leader emotional intelligence, transformational leadership, trust, and team commitment. The research framework of this study is based on those by MacKenzie et al. (2001), Schlechter and Strauss (2008), in which trust comprises trust in supervisors and trust within a team. The participants of this study are real estate brokers in Kaohsiung City. Structural equation modeling (SEM) was employed to explore the relationships and effects between leader emotional intelligence, trust, organizational commitment, transformational leadership, transactional leadership, and job performance.

Previous studies on the job performance of real estate brokers have mostly focused on how brokers' task performance, job satisfaction, and loyalty affected their job performance (e.g., Abelson et al., 1990; Wang and Netemeyer, 2002; Lee and Shen, 2008; Acharya et al., 2010; Lee et al., 2017). To the best of the authors' knowledge, few studies have collectively explored the effects of leader emotional intelligence, transformational leadership, and transactional leadership on job performance in the real estate brokerage industry. These three factors, however, have been examined for employees working in construction (Pryke et al., 2015), manufacturing (Schlechter and Strauss, 2008), and finance (MacKenzie et al., 2001; Walumbwa et al., 2008). Therefore, this study distinguishes itself by focusing on the real estate brokerage industry. The conceptual model in this study is established on three important constructs—leader emotional intelligence; transformational leadership; and transactional leadership, in conjunction with trust; organizational commitment; and job performance, whereby trust comprises trust in supervisors and trust within a team. Interestingly, this study employed three latent variables—leader emotional intelligence, transformational leadership, and transactional leadership, which are individually distinct in theoretical and definitional terms. The two leadership styles are conceptually different and are often included in model analysis [see Druskat (1994), Kalsoom et al. (2018), Purwanto et al. (2020)]; only one of these two mutually exclusive styles affects trust and job performance. The objectives of this study are as follows:

(1) To explore the effects of leader emotional intelligence, transformational leadership, transactional leadership, trust in supervisors, trust within a team, and organizational commitment on employee job performance.
(2) To examine whether leader emotional intelligence indirectly affects job performance through the mediator variables of trust within a team and transformational leadership.
(3) To examine whether transformational leadership and transactional leadership indirectly affect job performance through the mediator variable of trust in supervisors.

The participants of this study were real estate brokers in Kaohsiung City, Taiwan. According to the Real Estate Information Platform[1] of the Ministry of the Interior, as of January 31, 2021, there are 659 registered real estate brokerage companies, Of these companies/offices, 659 operate normally, employing a total of 4377 real estate salespeople and 749 brokers (227 of whom double as land administration agents). Kaohsiung City ranks fifth (after Taichung, Taipei, New Taipei, and Taoyuan cities, respectively) among all cities and counties in Taiwan in terms of the number of real estate offices and employees, evidence of its considerable real estate industry size. In those 659 offices, each has two supervisors—a manager and a vice manager, which means there are at least 1318 supervisors in Kaohsiung City's real estate brokerage industry. Furthermore, the remuneration scheme in the real estate brokerage industry differs from other industries as a broker's income is determined from their commissions (brokers are ranked by their income into high, intermediate, and common).[2] Moreover, because little prerequisite experience is required for newcomers, the brokerage industry has a particularly high turnover rate[3] (Ministry of Labor, 2020).

Interestingly, business models in the real estate brokerage industry itself vary by business model (direct sales or franchise). Specifically, the selection of managers (leaders), remuneration scheme, employee management, and working hours differ considerably between the two business models. However, the focal points of this study are the effects of leader emotional

[1]The Real Estate Information Platform of the Ministry of the Interior collates information on real estate appraisers, land administration agents, real estate brokerage, and rental and housing services. Website: https://resim.land.moi.gov.tw/Home/Queryindex. Last accessed on 2022.4.16.

[2]High-ranking brokers do not have a base salary and rely solely on sales incentives; intermediate brokers have a low base salary and high sales incentives; and common brokers have a high base salary and low sales incentives.

[3]Analysis on the Real Estate Industry, Ministry of Labor. Available at https://www.mol.gov.tw/media/1695/al01.pdf. Last accessed on 2020.2.20.

intelligence, leadership style, organization commitment, and trust on employee job performance. SEM can be used to examine the differences between the coefficients of business models, selection of managers (leaders), remuneration scheme, employee management, and working hours, such as whether the effects of trust on job performance differ by business model. However, these examinations were not performed in the present study due to the sheer complexity of the analyses as well as being out of the scope of study. In addition, even if business models in the real estate brokerage industry itself vary by business model, then the selection of managers (leaders), remuneration scheme, employee management, and working hours will vary considerably. Job performance is a latent variable in this study and consists of task performance and contextual performance. The former refers to challenging work tasks and demands, work efficiency, and overall efficiency, which may differ systematically and therefore have no consistency in terms of task performance-related demands. For the sake of prudence, this study analyzed the differences between direct sales and franchise operations through cluster analysis in SEM. The resulting chi-square was greater than that of the non-clustered data, and the chi-square test for difference revealed that the non-clustered data had a better fit. On this statistical basis, it can be concluded that brokers in direct sales offices do not differ significantly in terms of estimated coefficients.

LITERATURE REVIEW AND RESEARCH HYPOTHESES

Leader Emotional Intelligence, Transformational Leadership, Trust Within a Team, and Job Performance

Wong and Law (2002) suggested that emotionally intelligent and mature leaders exhibit higher awareness toward their own emotions as well as their subordinates' and adopt mentally supportive actions that positively affect their subordinates' job satisfaction and potential job performance. Pryke et al. (2015) found that in architectural team communications, project managers' emotional intelligence (including their emotional sensitivity and emotional performance) has important impacts on the communication between managers and their subordinates. A project manager with excellent leadership is also one of the factors that contributes to a project's success. Shahhosseini et al. (2012) studied the job performance of branch managers in the banking sector and found that the proper utilization of emotional intelligence increases the job performance and organizational effectiveness of managers and subordinates. Based on the above, this study proposes the following:

> H1: Leader emotional intelligence has a significant and positive effect on job performance.

Bass (1985) identified four prerequisite behaviors among transformational leaders: intellectual stimulation, individualized consideration, idealized or charismatic influence, and inspirational motivation. Sosik and Megerian (1999) found that emotional intelligence is a promoter of transformational leadership. For instance, emotionally intelligent leaders have the ability to empathize with their employees and thereby exhibit individualized consideration to employees to overcome their difficulties at work. Employees who perceive their leader to be emotionally competent know that the leader is equipped with transformational leadership skills. Leaders who are able to perceive their own emotional competence and comprehend their influence over their subordinates enhance the effectiveness of their leadership (Day and Carroll, 2004). Transformational leaders and emotionally intelligent leaders share highly similar traits and behaviors (Barling et al., 2000; Prati et al., 2003). Leaders who exhibit excellent self-control over their emotions become a mainstay for their followers and increase the trust and respect of their followers. Leaders who are able to recognize the strengths, weaknesses, and traits of an employee allocate suitable tasks that empower the employee to help attain set organizational goals (Harms and Credé, 2010). Based on the arguments above, this study proposes Hypothesis 2 as follows:

> H2: Leader emotional intelligence has a significant and positive effect on transformational leadership.

Podsakoff et al. (1990) proposed six key behaviors of transformational leaders: articulating a vision, providing a suitable work model for their subordinates, fostering the acceptance of team goals among employees, having high performance expectations, providing individualized support, and advocating innovation. MacKenzie et al. (2001) explored the leader behaviors of insurance salespeople and demonstrated that the individualized support provided by transactional leaders to their subordinates reflects the leaders' respect for their subordinates' opinions. These leaders also attend to their subordinates' feelings and needs and assist in the development of their careers, thereby positively and significantly affecting their job performance. Walumbwa et al. (2008) focused on bank employees in the American Midwest and found that transformational leadership positively and significantly affects employee job performance via the mediator variables of enhancing employees' self-efficacy and defining work goals. Liaw et al. (2010) showed that by enhancing employees' customer orientation, transformational leadership positively affects employees' service performance such that employees are willing to spend more time and effort to satisfy customers' needs and successfully establish long-term service relationships with them. Noruzy et al. (2013) used SEM to explore the relations between transformational leadership, organizational learning, knowledge management, organizational innovation, and organizational performance in the Iranian manufacturing industry. The results reveal that transformational leadership has a positive effect on organizational performance and creates social environments that benefit the organization, thereby prompting subordinates to increase their job performance by engaging in higher-level knowledge management-related activities (such as organizational learning, organization management, and organizational innovation). Based on the above, this study proposes Hypothesis 3 as follows:

H3: Transformational leadership has a significant and positive effect on job performance.

Prati et al. (2003) found that emotionally intelligent teams build a higher level of trust, and leader emotional intelligence influences employees' trust in their leader's leadership. Schlechter and Strauss (2008) argued that emotionally intelligent leaders resolve conflicts through constructive approaches and establish human-centered cooperative relationships. In short, leader emotional intelligence has a significant and positive effect on trust among team members. Chang et al. (2012) noted that by helping team members control their emotions, emotionally intelligent leaders can prevent the negative impacts of employees' overreactions and strengthen trust among team members. Liu et al. (2012) suggested that leader emotional intelligence, behaviors, and personalities are beneficial for fostering team morale and has positive effects on trust, communication, and participation among team members. Leaders also use their emotional intelligence to improve their trust in their team (Kauffmann and Carmi, 2014). On the basis of the above, this study proposes Hypothesis 4 as follows:

H4: Leader emotional intelligence has a significant and positive effect on trust in a team.

Transformational Leadership, Transactional Leadership, Trust in Supervisors, and Job Performance

Yang and Mossholder (2010) suggested that employees' trust in their leader is based on the leader's perceived degree of impartiality and sincerity. Trust is conceptualized as employees' belief in and loyalty to their leader. According to the study by MacKenzie et al. (2001), the core behaviors of a transformational leader (i.e., articulating a vision, providing a suitable work model for their subordinates, fostering the acceptance of team goals among employees, and providing individualized support) help employees believe that their leader will reward them based on their efforts. Therefore, transformational leadership has a positive and significant effect on trust in supervisors. Schlechter and Strauss (2008) highlighted that transformational leadership has a positive influence on trust in supervisors, because transformational leaders inspire their employees to achieve goals. Transformational leaders are also considerate of their employee's well-being and motivate them to perform better, thereby increasing their trustworthiness among their employees. Jung and Avolio (2000) agreed that transformational leadership has a significant and positive effect on trust in supervisors. The authors suggested that throughout the process of achieving organizational goals, transformational leaders will exhibit their firm beliefs and altruism in order to motivate their employees to achieve organizational goals. Hence, when employees often highly regard their leader, they identify with the leader's beliefs and have high respect toward them. Braun et al. (2013) studied the relationships between transformational leadership, trust within a team, trust in supervisors, and team performance among members in academic institutions. The results indicate that trust in supervisors mediates the positive and significant effect of transformational leadership on job satisfaction. Employees who trust their managers are more satisfied with their jobs because they are aware of the high level of concern and attention paid by their managers toward them (Yang and Mossholder, 2010). Weng and Li (2015) demonstrated that the behaviors of transformational leaders are manifested in their values, beliefs, and will. Team members who accept and internalize these traits naturally trust their supervisors. Asencio and Mujkic (2016) also found that transformational leadership has a significant and positive effect on trust in supervisors. Altunoğlu et al. (2019) found that a leader's transformational leadership style, as perceived by employees, generates a positive impact on affective trust. Islam et al. (2021) agreed that transformational leadership has a positive effect on trust. Therefore, this study proposes Hypothesis 5, as follows:

H5: Transformational leadership has a significant and positive effect on trust in supervisors.

Employees who develop trust in their leader identify with their organization, which motivates them to work harder and achieve a better performance (Schaubroeck et al., 2013). In contrast, employees who lose trust in their leader are more likely to feel discontent. As a result of such negativity, employees feel more burnt out and have less enthusiasm in their work tasks (Bechtoldt et al., 2007). Jung and Avolio (2000) found that trust in leaders has a positive and significant effect on job performance. Dirks and Ferrin (2002) also agreed that trust in supervisors has positive effects on employees' attitudes, behaviors, and job performance. Asencio and Mujkic (2016) highlighted that leaders should emphasize transformational leadership in order to strengthen interpersonal trust within their organization, motivate employees, and enhance organizational effectiveness. Mo and Shi (2017) studied employees working in a pharmaceutical retail company and found that trust in supervisors has a positive effect on employee job performance. Employees who distrust their leader lack the courage to communicate or express their feelings to them. Consequently, this results in a negative morale and climate within the organization. The literature above leads to the proposal of Hypothesis 6, as follows:

H6: Trust in supervisors has a significant and positive effect on job performance.

Nanjundeswaraswamy and Swamy (2014) observed that the relationship between transactional leaders and their subordinates is contract-based. Transactional leaders emphasize the needs and supervise the performance of their employees. They use reward and punishment schemes as a means for motivating employees to achieve organizational goals, i.e., employees are rewarded for their remarkable contributions to their company or punished if otherwise. Indeed, employees' performance and future rewards are dependent on their job performance (Bass, 1985). As proposed by Franco and Matos (2015), transactional leadership is more effective when an organization encounters uncomplicated and definitive problems. In times of crisis, transactional leadership within an organization assists employees to focus on completing tasks that help overcome the crisis.

Masa'deh et al. (2016) revealed that transactional leadership has significant and positive effects on job performance as it facilitates the enhancement of knowledge sharing within the organization and improves employees' job performance. In their study on Malaysian private pharmacies, Basri et al. (2017) demonstrated that transactional leadership has significant and positive effects on job performance. Based on the findings of previous studies, this study proposes Hypothesis 7 as follows:

> H7: Transactional leadership has a significant and positive effect on job performance.

Jung and Avolio (2000) indicated that transactional leaders gravitate toward gaining trust from their followers through contracts or exchanges of interests. Transactional leadership has a significant and positive effect on trust in supervisors. It also has an indirect and positive effect on job performance through the mediator variable of trust in supervisors. MacKenzie et al. (2001) highlighted how transactional leadership has a significant and positive effect on trust in supervisors. Transactional leaders who adopt contingent reward behaviors strengthen salespeople's trust in supervisors as they believe managers will allocate rewards based on their sales performance. Asencio and Mujkic (2016) analyzed the leadership behaviors and trust in supervisors among American federal government employees using multivariate regression. The results indicate that employees trust the beliefs and behaviors of leaders who adopt transactional leadership behaviors as the employees believe that these leaders are able to distinguish between rewards and punishments, i.e., high-performing employees are rewarded whereas low-performing employees or those who made errors are corrected or punished. Greenberg (2003) also noted that in organizational management, leaders who are transactional gain the trust of their employees. On the basis of the above findings above, this study proposes Hypothesis 8 as follows:

> H8: Transactional leadership has a significant and positive effect on trust in supervisors.

Trust in Supervisors, Trust Within a Team, Organizational Commitment, and Job Performance

Jong and Elfring (2010) opined that organizational members who trust their teammates can reduce the suspicion and operational uncertainty within the team. Mutual trust between team members is established through good interactions and teamwork, and results in higher job performance. In contrast, team members who lack trust in one another tend to avoid interactions or cooperation to protect themselves from the actions of their teammates (Mayer and Gavin, 2005). This behavior ultimately has a negative impact on their performance as a team. Drescher et al. (2014) indicated that team members who build trusting relationships with one another are willing to put in extra effort to help their colleagues. As trust continues to build within the team, so does cooperation, thereby enhancing the team's performance. Furthermore, mutual trust between team members reduces the time spent on supervising each other (Langfred, 2004). Ultimately, team members focus more on their work tasks (Serva et al., 2005). When more team members exhibit responsible and trustworthy behaviors, the team will invest more efforts in improving their workflow and enhancing their performance. Setiawan et al. (2016) argued that the attribute of trust positively affects an individual's job performance. Varshney and Varshney (2017) agreed that trust has a positive effect on job performance. Based on the above, this study proposes Hypothesis 9 as follows:

> H9: Trust within a team has a significant and positive effect on job performance.

Becker and Billings (1993) defined organizational commitment as the relative strength of an individual's identification with, and engagement in, a specific organization. Organizational commitment consists of employees' varying levels of commitment toward their supervisor, team, department, and organization as a whole. Cohen and Prusak (2001) demonstrated that employee job satisfaction and organizational commitment are established on the basis of trust. Schlechter and Strauss (2008) revealed that trust among team members has positive and significant effects on organizational commitment. Teams who build work relationships rooted in trust are able to strengthen their cooperation, reduce conflict, enhance organizational commitment, and decrease turnover intention. Yamaguchi (2013) agreed that trust within a team has significant and positive effects on employee job satisfaction and organizational commitment, and Chang et al. (2012) noted that trust has significant and positive effects on organizational commitment. Based on the above, this study proposes Hypothesis 10 as follows:

> H10: Trust within a team has a significant and positive effect on organizational commitment.

A study by Jaramillo et al. (2005) on salespeople and non-salespeople demonstrated a positive relationship between organizational commitment and job performance. The relationship between the organizational commitment and job performance of salespeople is stronger than that of non-salespeople. Rose et al. (2009) found that organizational learning has a positive effect on job performance through the mediator variable of organizational commitment. Yeh and Hong (2012) highlighted the positive and significant effect of organizational commitment on job performance. Employees with intentions to stay within their organization dedicate themselves to completing their work goals based on their agreement with the organization's values and goals. Therefore, leaders would give appropriate rewards when employees achieve their work goals. Supervisors who are able to achieve their own commitments lead to negativity among their subordinates, who become unwilling to put effort in their work and thereby lower their performance. Fu and Deshpande (2014) examined the relationships between caring climate, job satisfaction, organizational commitment, and job performance in a Chinese insurance company. The empirical results showed that organizational commitment had a significant and positive impact on job performance. Jamal (2011) investigated employees working at the Malaysian and Pakistani

bases of a multinational company and found that organizational commitment had a significant and positive impact on job performance. In addition, work stress affected job performance through the moderator variable of organizational commitment. Based on the above, this study proposes Hypothesis 11 as follows:

> H11: Organizational commitment has a significant and positive effect on job performance.

Lau and Moser (2008) highlighted the significant and positive effects of trust in supervisors on organizational commitment. Schlechter and Strauss (2008) agreed that trust in supervisors has a positive and significant effect on organizational commitment, and hence team leaders should develop organizational climates characterized by cooperation and trust. In so doing, leaders establish good organizational commitment and achieve effective teamwork. In their study on managers in the manufacturing and finance industries, Sholihin and Pike (2009) demonstrated the significant and positive effects of trust in supervisors on organizational commitment. Goh and Low (2013) examined the relationships between servant leadership, trust in leaders, and organizational commitment. The results indicate that trust in leaders has a positive and significant effect on organizational commitment. Trust in leaders is vital as it motivates employees to accept their leader's beliefs and therefore facilitates the establishment of mutual cooperation between employees and leaders. The findings above lead to the proposal of Hypothesis 12 as follows:

> H12: Trust in supervisors has a significant and positive effect on organizational commitment.

Regarding the mediating effects of the variables, briefly, this study also included several indirect and direct effects in its scope, including whether leader emotional intelligence indirectly affects job performance through the mediator variables of trust within a team and transformational leadership, as well as whether transformational leadership and transactional leadership indirectly affect job performance through the mediator variable of trust in supervisors.

RESEARCH DESIGN

The Research Framework

This research framework of this study, as shown in **Figure 1**, combines the three constructs of leader emotional intelligence, transformational leadership, and transactional leadership with a traditional job performance model that consists of trust within a team, trust in supervisors, organizational commitment, and job performance.

Operational Definitions of the Variables

Leader Emotional Intelligence

Emotional intelligence refers to an individual's ability to identify, express, understand, and evaluate their own emotions as well as those of others (Schlechter and Strauss, 2008). Davies et al. (1998) proposed a four-component model of emotional intelligence that consists of identifying one's own emotions, identifying the emotions of others, utilizing emotions, and managing one's own emotions. Identifying one's own emotions refers to a behavioral agent's observations and identification of their own emotions, which concerns their own subjective views of themselves. Since this study focuses on real estate brokers, the aforementioned component is excluded because how the emotional intelligence of leaders (store managers) affects their job performance as real estate brokers and how store manager perceive their own emotions do not fall within the scope of this study. This study refers to Davies et al.'s (1998) definition of emotional intelligence and suggests that the emotional intelligence of real estate brokers consists of the three dimensions of identifying others' emotions, utilizing emotions, and managing one's own emotions.

Transformational Leadership Style

Leaders with transformational leadership leverage their charm to motivate employees or meet their individual demands, thereby enhancing their job efficiency, confidence, and sense of authority, allowing them to surpass their own expectations and improve their job performance (Bass, 1985). Bass (1985), Bass and Avolio (1997) distinguished four traits of a transformational leader: idealized influence, inspirational motivation, intellectual stimulation, and individualized consideration. Podsakoff et al. (1990) proposed six key behaviors of transformational leaders: articulating a vision, assisting subordinates in finding suitable work models, fostering the acceptance of team goals among employees, having high performance expectations, providing individualized support, and advocating innovation. Following Bass' (1985), Bass and Avolio's (1997) definitions, the operational definition of transformational leadership in this study is a four-component model consisting of idealized influence, inspirational motivation, intellectual stimulation, and individualized consideration, whereby, in addition to evoking employees' work motivations, supervisors encourage and express concern toward their employees promptly to help them achieve their work goals. Idealized influence refers to the ability of store managers to articulate organizational visions and influence employees' job attitudes by establishing ambitious visions and objectives. Inspirational motivation refers to store managers' behaviors in evoking employees' job motivations, boosting their confidence, and enhancing their morale at work. Intellectual stimulation refers to store managers' capabilities to encourage workers to examine the difficulties and predicaments they encounter at work, while also providing guidance and assistance to stimulate their problem-solving skills. Individualized consideration refers to the individualized assistance and support provided by store managers, as well as their respect for the demands and feelings of their employees.

Transactional Leadership Style

In the concept of transactional leadership, the leader–employee relationship is established on the basis of transactions. Leaders set well-defined work goals and standards and implement rigorous reward or punishment systems so that employees understand that they will be rewarded in return for achieving the objectives of the organization (Bass, 1985; Jung and Avolio, 2000).

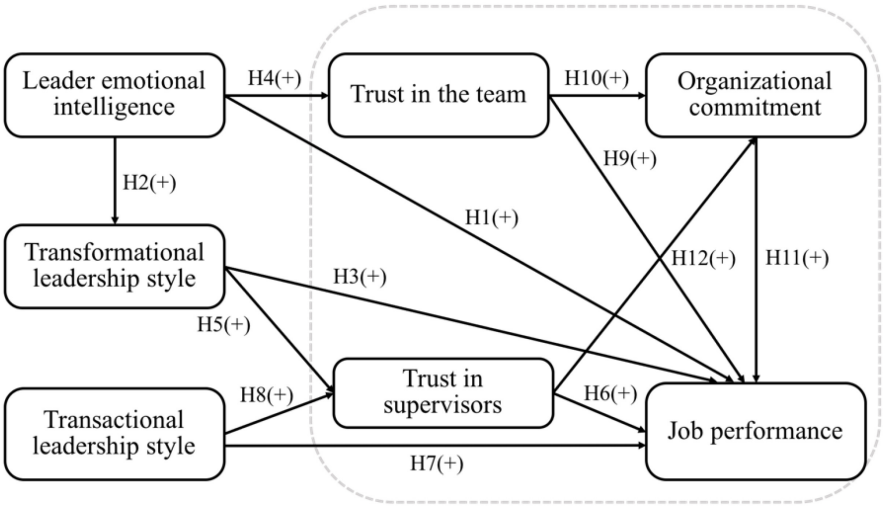


FIGURE 1 | The research framework proposed in this study.

Bass and Avolio (1997) examined transactional leadership through contingent reward and management by exception. The former refers to the provision of individualized and materialistic rewards, while the latter refers to active and passive management models. This study refers to Bass and Avolio's (1997) study and gives the operational definition of transactional leadership as leaders' utilization of contingent rewards and management-by-exception measures to set well-defined work goals and standards, implement systems of reward or punishment and enhance employees' job performance.

Trust Within a Team

Langfred (2004) explained that trust within a team is the aggregate perception of the level of trust among team members. With a high level of intra-team trust, team members focus more on their own tasks and less on monitoring one another. Staples and Webster (2008) defined trust within a team as the trusting relationships between coworkers and peers; that is, team members approve of each other's job competence, acknowledge the mutual trust between them, and treat one another fairly. Jong and Elfring (2010) concluded that intra-team trust is the trusting relationships team members establish amongst themselves based on their beliefs in one another's professionalism, dedication, work ethic, personality, and interpersonal relations. Based on the definition developed by Langfred (2004), Staples and Webster (2008), Jong and Elfring (2010), this study defines trust within a team as the level of trust between real estate brokers, and consisting of reliability, work ethic, impartiality, and interpersonal relations.

Trust in Supervisors

Lewis and Weigert (1985) explained trust in supervisors as employees believing in their supervisors' leadership capabilities, ascertaining that the decisions made by their supervisors are beneficial for organizational developments, and perceiving supervisors as honest and considerate individuals that care about their employees' demands and provide adequate assistance. MacKenzie et al. (2001) defined trust in supervisors as salespeople's belief in their leaders' impartiality and honesty. McAllister (1995) identified two types of trust: (1) Cognition-based trust, whereby a trustor perceives a trustee as trustworthy and competent based on their capabilities, expertise, or personalities (integrity, honesty, benevolence, and impartiality); (2) Affect-based trust, emphasizes the inclusion of emotional factors as the basis of trust between a trustor and a trustee, as well as the belief that the trustee will reciprocate these emotional efforts by showing the trustor their support, honesty, and care. This study adopts the cognitive theories of trust developed in the studies by McAllister (1995), MacKenzie et al. (2001) and defines trust in supervisors as the trusting relationship derived from a supervisor's personality traits such as integrity, honesty, benevolence, impartiality, and faithfulness.

Organizational Commitment

This study measured several sub-dimensions of organizational commitment to achieve a deeper understanding of its implications. Mowday et al. (1982) developed a model of organizational commitment that consists of three sub-dimensions: (1) Propensity to identify with organizational goals, in terms of which an employee identifies with their organization by accepting the organizational goals and values and developing cohesion; (2) Willingness to contribute to the organization, which refers to an employee's willingness to exert considerable effort for the organization; and (3) Tendency to retain in the organization, which refers to the desire to remain in the organization. Allen and Meyer (1990) developed an organizational commitment model with three components (affective, continuance, and normative commitment) as follows: (1) Affective commitment is generated from emotional attachment; (2) Continuance commitment is generated from perceiving the cost of leaving the organization; (3) Normative commitment is the sense of obligation to remain in the organization. This study

adopts Allen and Meyer's (1990) three-component model of organizational commitment to measure the commitment of real estate brokers to their branch offices.

Job Performance

Borman and Motowidlo (1993) described job performance as the aggregate of an individual's behaviors in relation to organizational goals, and is measured by their contribution to these goals. Job performance is analyzed through task performance and contextual performance. The former refers to the outcomes of an employee's expected or assigned tasks, quantitatively measured through indicators such as productivity and sales volume. The latter refers to the voluntary behaviors or performance of an employee, quantitatively expressed through the evaluations of supervisors and team members or the employee's self-perceptions. Byars and Rue (1994) defined job performance as the extent to which an employee's behavior contributes to organizational goals. The authors measured job performance based on the overall performance on the three dimensions of efficiency, effectiveness, and efficacy. Motowidlo and Van Scotter (1994) extended Borman and Motowidlo's (1993) classification and definition of job performance, describing task performance as employees' task outcomes measured by the extent to which they complete their organizational tasks while meeting the demands of their own tasks (job descriptions, standard operating procedures, and ad-hoc requests from supervisors). Contextual performance refers to an employee's ability to perform prescribed activities that are not officially part of their work, enthusiasm in completing tasks, willingness to cooperate with and assist others, compliance with organizational regulations and procedures at their own cost, and support and defense of organizational goals. These behaviors are voluntarily expressed by employees and cannot be coerced by the organization. This study adopts the two sub-dimensions of task performance and conceptual performance delineated in Motowidlo and Van Scotter's (1994) to measure the job performance of real estate brokers.

Questionnaire Design

The questionnaire in this study consists of two sections. The first section consists of items pertaining to leader emotional intelligence, transformational leadership, transactional leadership, trust within a team, trust in supervisors, organizational commitment, and job performance. The items concerning leader emotional intelligence are developed on the basis of the study by Davies et al. (1998) and comprise three aspects—identifying emotions, utilizing emotions, and managing one's own emotions. The items are revised according to the questionnaire by Law et al. (2004). Each aspect consists of two items, for a total of six items. The items concerning transformational leadership are developed on the basis of the studies by Bass (1985); Bass and Avolio (1997). There are four aspects—idealized influence, inspirational motivation, intellectual stimulation, and individualized consideration. The items are revised according to the questionnaire by Masa'deh et al. (2016). Each aspect consists of two items, for a total of eight items. The items concerning transactional leadership are developed on the basis of the study by Bass and Avolio (1997). There are two aspects—contingent reward and management by exception. The five items are revised according to the questionnaire by Masa'deh et al. (2016). The five items concerning trust within a team are designed according to the studies by Langfred (2004), Staples and Webster (2008), Jong and Elfring (2010). The five items concerning trust within a team are designed according to the studies by McAllister (1995), MacKenzie et al. (2001). Allen and Meyer (1990) classified organizational commitment into affective commitment, continuous commitment, and normative commitment. In this study, the items concerning organizational commitment cover these three aspects and are revised according to the studies by Meyer and Allen (1991), Fu and Deshpande (2014). Each aspect covers three items, for a total of nine items. The items concerning job performance are based on the study by Motowidlo and Van Scotter (1994), in which job performance consists of task performance and contextual performance. The questionnaire items are revised according to the studies by Fu and Deshpande (2014), Masa'deh et al. (2016). Each aspect consists of three items, for a total of six items. All the questionnaire items are measured on a five-point Likert scale (1 = strongly disagree; 2 = disagree; 3 = neutral; 4 = agree; 5 = strongly agree). The second section of the questionnaire covers the participants' basic information, including their age, gender, tenure in the real estate brokerage industry, and annual income. The questionnaire in this study was designed alongside undergraduate students, who also used it as part of their graduation project. The questionnaire items are presented in **Appendix 1**.

Sampling and Data Collection

The participants in this study consisted of real estate brokers employed at real estate franchises in Kaohsiung City. The real estate companies, with the number of branch offices in brackets, included Taiching Realty (59), Taiwan Real Estate (13), Yung-Ching Realty (43), H&B Housing (47), Chinatrust Real Estate Co. (19), Sinyi Realty (31), U-trust Housing (6), Pacific Realtor (2), ETWARM Real Estate Co., Ltd. (4), and Century 21 Real Estate (4). The geographical scope of research covered eight administrative districts in Kaohsiung City: Tsoying, Sanmin, Sanmin, Gushan, Lingya, Hsinhsing, Cianjin, Cianjhen, and Fongshan. These eight districts, which constituted the former Kaohsiung City (the former county-administered Kaohsiung City merged with Kaohsiung County to form a special municipality in 2010), were selected for their high density of real estate brokerage companies. The aforementioned 10 real estate brokerage companies listed a total of 224 branch offices on their websites. To include all the companies in the scope of research, each branch office was classified by location (administrative district) and then arranged by franchise. Subsequently, each branch office in each franchise was sampled, thereby ensuring that the sample covered all ten franchises in all eight districts. Ninety-eight branch offices were sampled in this study; 10 copies of the abovementioned questionnaire were delivered by the researchers in person to each branch office. Due to the nature of their work, the brokers had to be out of their offices at times, and hence the responses were collected three to five days after being administered. The survey period ran from May 1, 2018 to June 1, 2018. Of the 980 questionnaires administered, 411 were returned,

of which 388 were valid, indicating an effective response rate of 39.59%. The expected sample size must be considered during sampling as it affects the accuracy of the estimation results. Assuming a tolerable error (d) of 0.05 and a level of significance (α) of 10%, this study requires a sample size (n) of 271 with a confidence level of 90%. This demand is met as there were 388 valid responses.

DESCRIPTIVE STATISTICS OF THE SAMPLE

Basic Participant Data

Among the valid responses, men accounted for 57.2% (199 people) of the participants and women accounted for 42.8% (149 people) of the participants. The mean age of the participants was 41 years. The eldest participant was 70 years old, and the youngest was 22 years old. Regarding marital status, married participants accounted for 51.1% (178 people), unmarried participants accounted for 44.0% (153 people), and participants with other marital statuses accounted for 4.3% (15 people). In terms of education level, university-level participants (including 4- and 2-year programs) accounted for the highest proportion of participants at 46.6% (162 people), followed by participants with high school (vocational) education or below, who accounted for 27.3% (95 people); participants with specialized education, who accounted for 20.1% (70 people); and finally, participants with master's degrees and above, who accounted for 4.3% (15 people). In terms of job tenure, participants with 1 to 3 years of experience accounted for the highest proportion at 33.0% (115 people), followed by those with less than a year's experience as well as those with 4 to 6 years of experience, who both accounted for 19.0% (66 people). In terms of company positions, participants in agent positions (brokers, salespeople) accounted for the highest proportion at approximately 91.7% (319 people), followed by branch managers, who accounted for 6.9% (24 people). In terms of company business model, franchises were the dominant model, accounting for approximately 81.0% (282 branch offices), whereas direct sales operations accounted for approximately 19.0% (66 branch offices). In terms of average annual income in the last 3 years, a majority of the participants (26.1%, 91 people) had an average annual income of between NT$310,000 and NT$500,000; followed by those with less than NT$300,000 (24.7%, 86 people); and then by those with an average annual income of between NT$510,000 and NT$700,000 (21.3%, 74 people).

Data Processing

Before a questionnaire survey, it is important to consider the problem of non-response bias, which refers to any of the two following conditions which may result in an insufficiency of information obtained from the sample: (1) The responses are not collected in time or were returned only after the participants had to be reminded to do so. (2) There are missing data in the responses or the sample structure is too concentrated on a certain population or level. In addition, non-response bias also arises when representative samples for certain categories are missing in the questionnaire, which affects the completeness of the sample structure and thereby creates statistical bias (Chen and Wang, 2011). The questionnaire in this study was delivered by the researchers in person to each branch office, and the responses were collected 3 to 5 days later. Some of the responses were not returned in time, and the researchers had to remind the participants to return their responses. Therefore, the recovery process was completed twice, and a total of 348 valid responses were recovered. To ensure that the sample structure did not differ significantly between each recovery, as well as to ensure that the recovered data can be inferred to a population, the non-response bias test process proposed by Armstrong and Overton (1977) was employed to check for non-response bias in the sample. Firstly, the 162 responses returned in time were classified as Group 1, whereas the other 182 returned later were classified as Group 2. Next, the homogeneity test in the chi-square test (Armstrong and Overton, 1977) was used to check for homogeneity or consistency in the responses of both groups regarding the participants' basic data (gender, age, marital status, education level, tenure, position, business model, and mean annual income). The results of the chi-square test are summarized in **Table 1**. All the *p*-values are greater than 0.05, and the null hypothesis is not rejected. This shows that there is consistency in the basic participant data of both groups. In other words, the non-response bias in the questionnaire survey is not significant, and the recovered data can be inferred to the population.

Reliability and Validity Analysis

A reliability analysis checks the stability and reliability of a dataset, and is frequently measured using the Cronbach's α. According to DeVellis (2016), a Cronbach's α greater than 0.70 indicates that a scale has good consistency and stability. In this study, the Cronbach's α of the seven latent variables ranged from 0.912 to 0.973 and were all greater than 0,70. Hence, the questionnaire designed in this study has a remarkably high reliability level (see **Table 2**).

Validity refers to the extent to which the measured variables of a scale are able to accurately measure the theme of a study (Chen, 2005). Validity consists of content validity, convergent validity, and discriminant validity. With regard to the content validity, the questionnaire in this study was designed after referring to and revising the questionnaire items used by relevant scholars from home and abroad. The questionnaire had to be in line with the scope of study and the research motivations.

TABLE 1 | Results of the chi-square test for non-response bias.

Item	Chi-square statistic	Degree of freedom	*p*-Value
Gender	0.898	1	0.343
Age	8.348	5	0.138
Marital status	0.701	2	0.704
Education level	2.244	3	0.523
Tenure	6.836	8	0.554
Position	0.674	2	0.714
Business model	0.122	1	0.727
Mean annual income	3.526	9	0.940

TABLE 2 | Cronbach's α of each latent variable.

Variable	Cronbach's α
Leader emotional intelligence	0.936
Transformational leadership	0.973
Transactional leadership	0.927
Trust within a team	0.927
Trust in supervisors	0.956
Organizational commitment	0.912
Job performance	0.941

The items were developed, screened, and revised following the researchers' discussions with real estate professionals. Therefore, the questionnaire should have a considerable level of content validity, and the convergent validity and discriminant validity are discussed in a subsequent section.

EMPIRICAL RESULTS AND DISCUSSION

The empirical results of this study are presented in terms of the measurement model and the structural model.

Analysis of the Measurement Model

According to the recommendations of Nunnally and Bernstein (1994), DeVellis (2016), a construct has a high reliability if the Cronbach's α is greater than 0.70, whereas a Cronbach's α ranging from 0.50 to 0.70 is considered acceptable. As shown in **Table 3**, the Cronbach's α of each variable is greater than 0.5, suggesting that the measured variables in this study have remarkable internal consistency and that the questionnaire results are within an acceptable range.

This study employed SEM and used Anderson and Gerbing's (1988) two-step approach for analysis.

Construct validity is measured in terms of factor loadings in this study. As shown in **Table 3**, the factor loading of each measured variable is statistically significant, demonstrating their excellent convergent validity. In addition, confirmatory factor analysis was performed to inspect the convergent validity and the discriminant validity of each construct. Regarding the evaluation of convergent validity, Anderson and Gerbing (1988) recommended that the measurement model of the structural model can be used to determine whether each measured variable can suitably measure each latent variable. As shown in **Table 4**, the factor loading of each measured variable is greater than 0.7 and statistically significant. Taken together, the questionnaire has an excellent convergent validity.

With regard to discriminant validity, Fornell and Larcker (1981) suggested that for a construct to discriminate well with other constructs, the square root of the average variance extracted (AVE) of a specified latent variable must be higher than the correlation coefficients of the other latent variables. For example, the correlation coefficient between leader emotional intelligence and transformational leadership is 0.873, which is smaller than the square root of their respective AVE (0.903 for leader emotional intelligence and 0.961 for transformational leadership). Therefore, based on analogous deductions, there is considerable discriminant validity between the constructs, as shown in **Table 3**.

Analysis of the Structural Model

Evaluation of the Theoretical Model

The fit of the theoretical model was first tested using the chi-square statistic χ^2(p-Value). If the value is statistically significant, then the theoretical model is not consistent with the data distribution structure, which means that other indices must be used instead. Common indices include the ratio of the chi-square statistic to the degree of freedom (χ^2/df), the goodness of fit index (GFI), the root mean square residual (RMR), the root mean square error of approximation (RMSEA), the adjusted goodness of hit index (AGFI), the normal fit index (NFI), and the comparative fit index (CFI). Bagozzi and Yi (1988) proposed three approaches for measuring the fit of a model, i.e., preliminary fit criteria, fit of internal structure of the model, and overall model fit. Each approach is described as follows:

(1) Preliminary fit criteria:

According to **Table 4**, the factor loadings of the seven latent variables were all statistically significant and were greater than 0.7. There were also no negative values in the measured error variances. In general, the model of this study met the preliminary fit criteria.

(2) Fit of the internal structure of the model:

According to Bagozzi and Yi (1988), there are three approaches for evaluating the fit of the internal structure of a

TABLE 3 | Correlation matrix of latent variables.

	Leader emotional intelligence	Transformational leadership	Transactional leadership	Trust within a team	Trust in supervisors	Organizational commitment	Job performance
Leader emotional intelligence	0.903						
Transformational leadership	0.873	0.961					
Transactional leadership	0.849	0.741	0.948				
Trust within a team	0.830	0.725	0.705	0.911			
Trust in supervisors	0.788	0.780	0.818	0.654	0.950		
Organizational commitment	0.569	0.534	0.544	0.565	0.604	0.908	
Job performance	0.628	0.619	0.598	0.666	0.604	0.674	0.951

The diagonal elements shown in this matrix are the square roots of the constructs' AVE.

TABLE 4 | Analysis of the questionnaire's reliability, factor loading, and average variance extracted.

Variable	Factor loading (unstandardized)	Factor loading (standardized)	Error variance	Reliability of measured variable	Composite reliability (CR)	Average variance extracted (AVE)	Structural equation assessment R^2
Leader emotional intelligence					0.930	0.816	
Identifying the emotions of others	1.000	0.839	0.170	0.704			
Utilizing emotions	0.980***	0.888***	0.105	0.789			
Managing one's own emotions	0.924***	0.794***	0.203	0.631			
Transformational leadership					0.980	0.923	0.762
Idealized influence	0.974***	0.911***	0.087	0.830			
Inspirational motivation	1.013***	0.954***	0.045	0.911			
Intellectual stimulation	1.044***	0.927***	0.080	0.860			
Individualized consideration	1.000	0.924	0.077	0.853			
Transactional leadership					0.946	0.898	
Contingent reward	1.018***	0.911***	0.078	0.830			
Management by exception	1.000	0.883	0.105	0.779			
Trust within a team					0.961	0.830	0.689
Trust within a team 1	1.000	0.863	0.143	0.745			
Trust within a team 2	0.991***	0.855***	0.151	0.730			
Trust within a team 3	0.952***	0.846***	0.150	0.716			
Trust within a team 4	0.975***	0.886***	0.109	0.785			
Trust within a team 5	0.832***	0.784***	0.182	0.614			
Trust in supervisors					0.979	0.903	0.736
Trust in supervisors 1	0.992***	0.858***	0.138	0.736			
Trust in supervisors 2	0.937***	0.917***	0.065	0.840			
Trust in supervisors 3	0.986***	0.896***	0.093	0.803			
Trust in supervisors 4	1.002***	0.906***	0.086	0.821			
Trust in supervisors 5	1.000	0.935	0.056	0.874			
Organizational commitment					0.933	0.825	0.415
Affective commitment	1.000	0.700	0.211	0.490			
Continuous commitment	1.602***	0.924***	0.089	0.854			
Normative commitment	1.690***	0.890***	0.152	0.792			
Job performance					0.950	0.905	0.596
Task performance	0.932***	0.931***	0.054	0.867			
Contextual performance	1.000	0.879	0.119	0.773			

** denotes $p < 0.1$, ** denotes $p < 0.05$, *** denotes $p < 0.01$.*

model, as follows: (1) The individual item reliability of each item is used to assess the construct reliability of measured variables to their corresponding latent variables, to validate whether a factor loading is greater than 0.5 and check the statistical significance of each loading. (2) The composite reliability (CR) of a latent variable is an aggregate of the reliabilities of its measured variables. It is used to check the consistency between the latent variables measured within a construct. Fornell and Larcker (1981) suggested that a CR greater than 0.6 indicates good reliability; the greater the CR, the higher the consistency of the internal construct indicators. The CRs of the latent variables in this study ranged were all greater than 0.9, reflecting the high consistency between the latent variables. (3) The AVE of a latent variable is the average percentage of variance of a measured variable that is explained by the latent variable (Chen, 2005). A high AVE indicates that the latent variable has a high convergent validity and reliability. Fornell and Larcker (1981) suggested that an AVE greater than 0.5 indicates good reliability. In this study, as shown in **Table 4**, the AVE values of the latent variables were all greater than 0.8, which indicates that the internal consistency of the questionnaire was acceptable.

(3) Overall fit of the model:

Bagozzi and Yi (1988) stressed that the fit of a structural model cannot be determined through a single indicator. Instead, the test results of the overall model reflect the fit of the structural model. Hair et al. (1998) noted that there are three types of overall model fit measures: absolute fit measures, incremental fit measures, and parsimonious fit measures. (1) Absolute fit measures determine the level of predictive covariates in the overall model. As shown in **Table 5**, the chi-square statistic (χ^2) was 957.692 ($p < 0.001$) and was statistically significant. The null hypothesis is thus rejected, as the assumption model in this study did not have a good fit with the observed model. The chi-square statistic is very sensitive to sample sizes; a large sample size will increase the chi-square statistic, which results in the rejection of the null hypothesis. To

TABLE 5 | Goodness-of-fit indices of the conceptual framework model developed in this study.

Statistic		Benchmark for an ideal fit	Results
χ^2 (*p*-value)		Not statistically significant	957.692 (0.001)
Absolute fit measures	χ^2/df	Smaller than 5	4.007
	GFI	Greater than 0.90	0.818
	RMR	Favorably smaller	0.041
	RMSEA	Favorably smaller, ideal if smaller than 0.05	0.093
Incremental fit measures	*AGFI*	Greater than 0.90	0.771
	NFI	Greater than 0.90	0.900
	CFI	Greater than 0.90	0.922
Parsimonious fit measures	*PNFI*	Greater than 0.50	0.779
	PGFI	Greater than 0.50	0.651

resolve this issue, it is necessary to consider other fit indices for evaluating the fit of the overall model (Chiu, 2006). The values of the other indices are as follows: the normal chi-square statistic (χ^2/df) was 4.007, which is smaller than 5; the RMR was 0.041; and the RMSEA was 0.093, which are both within an acceptable range; whereas the GFI was 0.818, which is close to 0.9. Generally speaking, the fit of the conceptual framework model is acceptable. (2) With regard to incremental fit measures, the AGFI was 0.771, which is close to 0.90; and the CFI was 0.922, which exceeds 0.90. Therefore, the incremental fit measures of the model are either acceptable or approximately acceptable. (3) Parsimonious fit measures, which are measured by adjusting the fit, determine the goodness of fit that can be obtained by each parameter estimate. The PNFI and PGFI in this study were 0.779 and 0.651, respectively, and were both greater than 0.50. This means that the parsimonious fit measures of the conceptual framework model devised in this study had a good fit. Furthermore, the values of all three fit measures attest to the good fit of the overall model.

Analysis of the Linear Structural Equation Model

(1) Empirical results:

The empirical results of the linear structural equation model are presented in **Figure 2** and **Table 6** alongside the corresponding standardized estimated coefficients.

According to the results, the estimated coefficient of leader emotional intelligence on job performance was -0.218 but failed to reach the 10% significance level. This suggests that real estate brokers' perceptions of their leaders' emotional intelligence had no significant effect on their job performance; H1 is thus not supported. The estimated coefficient of leader emotional intelligence on transformational leadership was 0.873 with a 1% significance level; therefore, H2 is supported. The estimated coefficient of transformational leadership on job performance was 0.231 with a 5% significance level. This suggests that real estate brokers' perceptions of their leaders' transformational leadership had a significant and positive effect on their job performance; H3 is thus supported. The estimated coefficient of leader emotional intelligence on trust within a team was 0.830 with a 1% significance level, thereby demonstrating the positive and significant effect of leader emotional intelligence on trust within a team; therefore, H4 is supported. The estimated coefficient of transformational leadership on trust in supervisors was 0.348 with a 1% significance level, thereby supporting H5. The estimated coefficient of trust in supervisors on job performance was 0.019 but failed to reach the 10% significance level. This demonstrates that employees' high level of perceived trust in their supervisor had no significant and positive effect on their job performance; hence, H6 is not supported.

The estimated coefficient of transactional leadership on job performance was 0.137 but failed to reach the 10% significance level. This demonstrates that employees' perceptions of their leader's transactional leadership behaviors had no significant and positive effect on their job performance; therefore, H7 is not supported. The estimated coefficient of transactional leadership on trust in supervisors was 0.534 with a 1% significance level. This suggests that employees who perceive that their leader exhibits the strong qualities of a transactional leader trust their leader even more; therefore, H8 is supported. The estimated coefficient of within-team trust on job performance was 0.351 with a 1% significance level. This shows that trust within a team has a significant and positive effect on job performance; hence, H9 is supported. The estimated coefficient of within-team trust on organizational commitment was 0.297 with a 1% significance level. This suggests that high levels of trust within a team strengthens employees' perception of organizational commitment; H10 is thus supported. The estimated coefficient of organizational commitment on job performance was 0.390 with a 1% significance level; H11 is thus supported. The estimated coefficient of trust in supervisors on organizational commitment was 0.409 with a 1% significance level; therefore, H12 is supported.

The empirical results indicate that leader emotional intelligence had no significant and direct effects on job performance but had indirect effects through the mediator variables of within-team trust and transformational leadership. Transformational leadership had direct effects on job performance but had no indirect effects through the mediator variable of trust in supervisors. Finally, transactional leadership had no direct or indirect effects (through trust in supervisors) on job performance.

(2) Discussion:

This suggests that real estate brokers' perceptions of their leaders' emotional intelligence had no significant effects on their job performance, and thus H1 is not supported. An examination by Sy et al. (2006) on foodservice employees revealed that the estimated coefficient of leader emotional intelligence on job performance was 0.2, but this failed to reach the 5% significance level. The authors attributed this result to certain tasks, such as cooking, cleaning, and equipment maintenance, involving limited interactions with others. Consequently, there were no significant effects of leader emotional intelligence on job performance. Because half the participants in the study sample

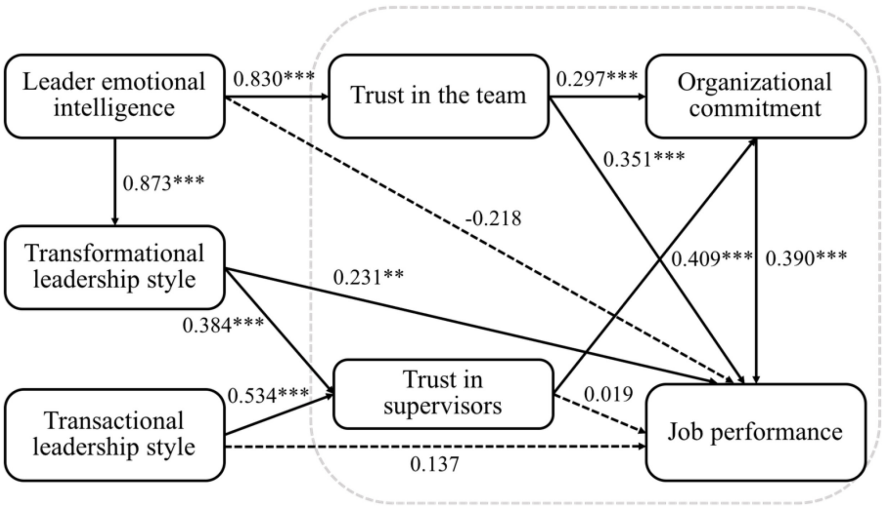


FIGURE 2 | The linear structural equation model (with standardized coefficients). * denotes $p < 0.1$, ** denotes $p < 0.05$, *** denotes $p < 0.01$.

TABLE 6 | Estimation results of linear structural equation modelling.

Hypothesis	Relationship between variables	Estimated coefficient	Standard error	t-statistic	*p*-Value	Outcome
H1	Leader emotional intelligence → Job performance	−0.218	0.181	−1.201	0.230	Not supported
H2	Leader emotional intelligence → Transformational leadership	0.873***	0.049	18.539	0.001***	Supported
H3	Transformational leadership → Job performance	0.231**	0.102	2.147	0.032**	Supported
H4	Leader emotional intelligence → Trust within a team	0.830***	0.053	15.918	0.001***	Supported
H5	Transformational leadership → Trust in supervisors	0.384***	0.050	7.165	0.001***	Supported
H6	Trust in supervisors → Job performance	0.019	0.092	0.212	0.832	Not supported
H7	Transactional leadership → Job performance	0.137	0.120	1.194	0.233	Not supported
H8	Transactional leadership → Trust in supervisors	0.534***	0.058	9.397	0.001***	Supported
H9	Trust within a team → Job performance	0.351***	0.089	3.874	0.001***	Supported
H10	Trust within a team → Organizational commitment	0.297***	0.046	4.486	0.001***	Supported
H11	Organizational commitment → Job performance	0.390***	0.086	6.403	0.001***	Supported
H12	Trust in supervisors → Organizational commitment	0.409***	0.048	6.088	0.001***	Supported

** denotes $p < 0.1$, ** denotes $p < 0.05$, *** denotes $p < 0.01$. All the estimated coefficients are standardized.*

were newcomers with less work experience, they required more time to become acquainted with their leaders. Only when they had enjoyed a long professional relationship would subordinates perceive their leaders' emotional intelligence. As such, this finding is similar to that of Sy et al. (2006). Wong and Law (2002) also found that leader emotional intelligence had no significant effects on job performance, which could be due to the participants' profession as civil servants. Specifically, the requirement of civil service exams and the bureaucratic context could distort the performance ratings of subordinates. As a whole, the results of this study do not support H1, but H2 is supported. This finding is in agreement with the findings of Prati et al. (2003), Schlechter and Strauss (2008). Indeed, leader emotional intelligence has a positive and significant effect on transformational leadership. Sosik and Megerian (1999) observed that emotional intelligence promotes transformational leadership behavior. Emotionally intelligent leaders have the ability to empathize with their employees and exhibit individualized caring behaviors that help employees overcome their difficulties at work. Employees who perceive their leader to be emotionally competent know that leader is equipped with transformational leadership skills.

H3 is supported. Bacha (2014) explored the relationships between transformational leadership, job performance, and job characteristics among employees in the French service and manufacturing industries. The results demonstrate the significant and positive effect of transformational leadership on job performance. Manaf and Latif (2014) showed that through the mediating effect of the cultural trait of adaptability, transformational leadership had a significant and positive effect on job performance. Transformational leadership is rooted in having a clear vision and stimulating motivation. Transformational leaders are sometimes able to give pragmatic assistance to employees to boost sales, therefore generating significant effects on employee job performance.

H4 is supported. Employees' strong perceptions of their leader's emotional intelligence increase the level of trust within their team. This finding is in line with Liu et al. (2012),

who found that emotionally intelligent leaders create positive team morale and build trusting relationships between members. When conflicts arise between team members, such leaders take appropriate measures to help them contain their emotions, resolve internal team conflicts, reduce the negative impacts of these conflicts, and enhance trust within the team (Chang et al., 2012). H5 is supported. This agrees with Asencio and Mujkic (2016), who found that transformational leadership has a positive and significant effect on trust in supervisors. Real estate brokers who perceive that their leader displays strong qualities of transformational leadership trust them even more. Schlechter and Strauss (2008) stated that transformational leaders help solve subordinates' problems and recognize their efforts. In turn, subordinates trust in their supervisors is strengthened, and they are willing to dedicate themselves to the supervisor or the company. In the process of achieving organizational goals, transformational leaders display their firm beliefs and altruism as a means of motivating their employees to achieve organizational goals. Hence, when employees often regard their leader highly, they identify with that leader's beliefs and have high respect for them (Jung and Avolio, 2000). In addition, transformational leaders are inspirational in the sense that they motivate employees to achieve goals while expressing concern for their well-being. Consequently, transformational leaders are perceived as trustworthy by their employees (Dirks and Ferrin, 2002). The core behaviors of a transformational leader, such as articulating a vision, providing a suitable work model for subordinates, fostering employees' acceptance of team goals, and providing individualized support, encourage employees to believe that their leaders will reward their efforts and performance accordingly (MacKenzie et al., 2001). Weng and Li (2015) find that the behaviors of transformational leaders are reflected in their values, beliefs, and will. When team members accept and internalize these traits they tend to trust their supervisors. Taken together, transformational leadership has a positive and significant effect on trust in supervisors (Asencio and Mujkic, 2016).

H6 is not supported. Sheng et al. (2005) argued that interpersonal trust has negligible effects on employee job performance. Instead, this performance could be affected by other factors, including communication between members' leadership styles, members' task performance, and type of teamwork. Because half the study participants were newcomers with limited work experience, they required more time to become acquainted and build trust with their leaders. This scenario could explain the insignificant impact of trust in supervisors on employee job performance. The empirical results, however, show that even though trust in supervisors had no direct effects on employee job performance, there was an indirect effect through organizational commitment. In other words, brokers display their trust in their supervisor through organizational commitment. Even though broker's trust in their supervisor did not have a direct and significant effect on their job performance, an indirect effect was generated through the mediator variable of organizational commitment.

H7 is not supported. Jung and Avolio (2000) indicated that whereas transactional leadership does not have a positive, direct, and significant effect on job performance, it does have a positive, indirect, and significant effect on job performance through the mediator variables of follower trust and value congruence. Shih (2006) concluded that transactional leadership had no significant effects on job performance because the remuneration scheme and standards of bank employees are fixed and well-defined. In this case, transactional leaders are unable to leverage their transactional traits to attain motivational effects. As the participants in this study are from the real estate brokerage industry, which has a defined set of duties and a clear remuneration scheme, transactional leaders may find their leadership skills have limited effectiveness within this context. H8 is supported. Asencio and Mujkic (2016) found that transactional leadership has a positive and significant effect on trust in supervisors. Transactional leaders are more likely to gain acceptance from their employees for their beliefs as the employees believe that such leaders are able to distinguish between rewards and punishments. In this regard, employees believe that they will receive the remuneration they deserve when they accomplish the goals set by their supervisor. Therefore, H8 is supported.

H9 is supported. Team members who trust their teammates can reduce suspicion and operational uncertainty within the team (Jong and Elfring, 2010). This mutual trust reduces the time members spend supervising one another (Langfred, 2004). Subsequently, team members focus more on their work tasks (Serva et al., 2005). Team members are willing to put in greater efforts to help their colleagues. As trust continues to build within the team, so does cooperation, thereby enhancing team performance (Drescher et al., 2014). Conversely, team members who distrust one another tend to avoid interaction or cooperation to protect themselves from the actions of their teammates (Mayer and Gavin, 2005). In real estate brokerage, a single customer may encounter different brokers throughout the different stages or forms of selling, such as product introduction, commissioned selling, and actual sale. Therefore, trust in their team is crucial for real estate brokers. Brokers who display responsible and trustworthy behavior during these different stages collectively improve their job performance. H10 is supported. This finding agrees with Yamaguchi (2013). Schlechter and Strauss (2008) argued that trust among team members contributes to better teamwork and organizational commitment. Geringer and Frayne (1993) likened the aspects of organizational commitment to construction materials; trust within a team is the concrete that firmly binds a brick wall. Strengthening trust within a team increases organizational commitment.

H11 is supported, and organizational commitment has a significant and positive effect on job performance. The results of this study are in line with Yeh and Hong (2012), Fu and Deshpande (2014). Indeed, organizational commitment has a significant and positive effect on job performance. Yeh and Hong (2012) concluded that employees who identify with their organization's goals and values are willing to retain their positions and dedicate themselves to the organization.

H12 is supported. This finding is in line with that of Goh and Low (2013); that is, trust in supervisors has a significant and positive effect on organizational commitment. Schlechter and Strauss (2008) commented that leaders of organizations should establish team environments characterized

by mutual trust and cooperation to enhance employees' organizational commitment and loyalty and thus achieve effective team operations. MacKenzie et al. (2001) reported that trust in supervisors is the overall feeling of dependence displayed by employees toward their organizational leader. Organizational commitment is a series of behavioral expressions generated by organizational members on the basis of trust in their organization (Meyer and Allen, 1991). Employees who trust their organization have relatively higher organizational commitment. When employees have higher trust in their managers, organizational commitment is beneficial to the company's sustainability (Sholihin and Pike, 2009).

CONCLUSION AND RECOMMENDATIONS

Theoretical Implications

Focusing on real estate brokers and their leaders (store managers), this study explored the effects of leader emotional intelligence, transformational leadership, transactional leadership, trust, and organizational commitment on the brokers' job performance. Leader emotional intelligence, transformational leadership, and transactional leadership constituted a traditional job performance model and served as important determinants of the job performance of real estate brokers. Trust in this study consisted of trust in supervisors and trust within a team, and the effects that both variables confer on job performance were also explored. Trust, commonly classified as individual trust of trust within a team, are both discussed in this study. In addition, leader emotional intelligence serves as an antecedent of trust within a team, whereas transformational leadership and transactional leadership serve as antecedents of trust in supervisors. In this regard, it is crucial to consider leader emotional intelligence, transformational leadership, and transactional leadership within analyzing trust in supervisors and trust within a team. Although previous studies have examined the relationships between leadership styles, emotional intelligence, and job performance individually, there is still a lack of comprehensive analyses on this topic. MacKenzie et al. (2001), Harms and Credé (2010), Bacha (2014), examined the effects of leadership style on job performance. Shahhosseini et al. (2012) examined the effects of emotional intelligence on job performance. Barling et al. (2000), Harms and Credé (2010) examined the association between leadership style and emotional intelligence. The lack of comprehensive analysis is addressed by this study examination of the relationships between leadership styles, emotional intelligence, and job performance by developing an integrated model of these variables on the basis of previous studies.

The empirical results indicate that leader emotional intelligence has positive and significant effects on transformational leadership and trust within a team, and transformational leadership has a positive and significant effect on trust in supervisors. In addition, transactional leadership has a positive and significant effect on trust in supervisors, and transformational leadership and trust within a team have positive and significant effects on employee job performance. Furthermore, trust within a team has a positive and significant effect on organizational commitment. However, leader emotional intelligence has no direct, positive, and significant effects on employee job performance, and trust in supervisors has no positive and significant effects on employee job performance. Moreover, transactional leadership also has no positive and significant effects on employee job performance. Even though the effects of leader emotional intelligence, trust in leaders, and transactional leadership on job performance do not support the hypotheses, the findings are consistent with real-life industry practices. This is due to the intense competition in the real estate industry, as it is easy for an individual to become a broker (Sinyi Realty, 2012).[4] Most brokers earn their wages based on their sales performance, and prioritize the resources provided by their company (leader) to assist them in boosting their sales performance, while internal management-related factors such as leader emotional intelligence are secondary in importance.

Collectively speaking, even though leader emotional intelligence and transactional leadership have no direct, positive, and significant effects on employee job performance, trust in supervisors or the trust within a team as shaped by the leadership styles and emotional intelligence of leaders will boost a team's spirit and morale, thereby affecting their job performance in an indirect fashion.

Practical Implications

Firstly, leader emotional intelligence affects employee job performance through trust within a team, organizational commitment, and transformational leadership. In other words, emotionally intelligent leaders create a trustworthy and harmonious climate within a team, which helps improve subordinates' job performance. Leader emotional intelligence includes leaders' ability to control their own emotions, as well as recognizing and managing the emotional conflicts of others. Additionally, D'Errico (2019) studied the effects that arise from the emotional intelligence displayed (happiness, calmness, sadness, and anger) by political leaders when discussing ethical issues such as immigration. The results showed that humble politicians often elicit humility in their facial expressions and evaluations when communicating with others. This induces a state of personal distress toward immigrants among voters, especially when the politician displays a sad facial expression. Therefore, leader emotional intelligence does not imply a leader's propensity to appease the public, but rather understanding their status and modifying strategies when appropriate.

Therefore, when training managers, real estate brokerage companies should stress the following actions: (1) Improving the observational skills of managers so that they can understand the needs of employees and customers, handle interpersonal conflicts effectively, and create harmonious and trusting work environments. (2) Strengthening the managers' ability to control their emotions and handle affairs in a rational manner. (3) Enhancing the leaders' emotional intelligence-related

[4]Source: Sinyi Realty (2012). Analysis of long-term investments. Website: https://statementdog.com/blog/archives/3225, Last accessed: 2022.3.30.

knowledge, which will help them resolve interpersonal conflicts in employees or customers.

Furthermore, transformational leadership has a direct effect on job performance and also has indirect effects through trust in supervisors and organizational commitment. When training managers, real estate brokerage companies should emphasize the development of leadership qualities in store managers so that they clearly understand their work goals and organizational visions, give timely motivations to employees, encourage innovation and self-improvement, take note of the opinions of their subordinates, satisfy employees' demands, and provide necessary assistance. Moreover, leaders should actively build reliable work climates in order to boost employees' loyalty and expectations toward their organization, thus increasing their willingness to dedicate themselves to the organization.

Lastly, transactional leadership affects job performance through trust in supervisors and organizational commitment. When training managers, real estate brokerage companies should emphasize the executive power of leaders by establishing well-defined work goals and content and appropriately implementing reward/punishment schemes. These measures will boost the discipline of team members as well as the authority of leaders. In addition, leaders should boost their integrity so that employees have faith and reliance in their organization, which generates a sense of belongingness to their organization. In this sense, they will believe that they will be rewarded for dedicating themselves to their job, thereby enhancing the employees' performance and the operations of the organization.

Limitations and Recommendations for Further Study

The questionnaire in this study was administered to real estate brokers, who completed all the items. It is suggested that the data could be obtained through more sources, such as the self-reported values of store managers and professionals. This enhances the understanding of the effects of leadership traits and styles on job performance. Next, further examination should be performed on the business models of the stores as well as the attributes of supervisors and employees alike. Finally, future studies can explore how other variables such as charismatic leadership, paternalistic leadership, knowledge sharing, personality traits, and corporate welfare affect the job performance of real estate brokers. Doing so would enhance the diversity and depth of the study.

In this study, SEM was used for analysis, which suffers from certain limitations, such as only being able to process the relationship between single-level variables, thus neglecting the effects of different variables. To overcome this issue, hierarchical linear modeling can be used for analysis, whereby variables are divided into the individual and aggregate levels, thus preventing the potential bias and underestimation that may arise in traditional statistical approaches (Courgeau, 2003). Going forward, future research can extend the factors in this study to analyze real estate properties at home and abroad. By examining the differences between various cultural backgrounds and environments (such as political environments), more information can be provided to facilitate the development and management of real estate business strategies.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct, and intellectual contribution to the work, and approved it for publication.

REFERENCES

Abelson, M., Kacmar, M., and Jackofsky, E. (1990). Factors influencing real estate brokerage sales staff performance. *J. Real Estate Res.* 5, 265–275. doi: 10.1080/10835547.1990.12090621

Acharya, R., Kagan, A., and Zimmerman, T. (2010). Real estate information technology: Influence of email marketing on real estate agent performance. *J. Real Estate Lit.* 18, 329–343. doi: 10.1080/10835547.2010.12090270

Allen, N. J., and Meyer, J. P. (1996). Affective, continuance, and normative commitment to the organization: an examination of construct validity. *J. Vocat. Behav.* 49, 252–276. doi: 10.1006/jvbe.1996.0043

Allen, N. J., and Meyer, J. P. (1990). The measurement and antecedents of affective, continuance and normative commitment to the organization. *J. Occup. Psychol.* 63, 1–18. doi: 10.1111/j.2044-8325.1990.tb00506.x

Altunoğlu, A. E., Şahin, F., and Babacan, S. (2019). Transformational leadership, trust, and follower outcomes: a moderated mediation model. *Manage. Res. Rev.* 42, 370–390. doi: 10.1108/mrr-01-2018-0036

Anderson, J. C., and Gerbing, D. W. (1988). Structural equation modeling in practice: a review and recommended two-step approach. *Psychol. Bull.* 103, 411–423. doi: 10.1037/0033-2909.103.3.411

Armstrong, J. S., and Overton, T. S. (1977). Estimating nonresponse bias in mail surveys. *J. Mark. Res.* 14, 396–402. doi: 10.2307/3150783

Asencio, H., and Mujkic, E. (2016). Leadership behaviors and trust in leaders: evidence from the US federal government. *Public Adm. Q.* 40, 156–179.

Bacha, E. (2014). The relationship between transformational leadership, task performance and job characteristics. *J. Manage. Dev.* 33, 410–420. doi: 10.1108/jmd-02-2013-0025

Bagozzi, R. P., and Yi, Y. (1988). On the evaluation of structural equation models. *J. Acad. Mark. Sci.* 16, 74–94.

Barling, J., Slater, F., and Kelloway, K. E. (2000). Transformational leadership and emotional intelligence: an exploratory study. *Leadersh. Organ. Dev. J.* 21, 157–161. doi: 10.1108/01437730010325040

Basri, H. H., Rashid, I. M. A., Abashah, A. B., and Samah, I. H. A. (2017). The impact of Malay transactional leadership style on employee performance: the Malaysia' pharmacy industry perspective. *Int. J. Inf. Bus. Manage.* 9:45.

Bass, B. M. (1985). *Leadership and Performance Beyond Expectations*. New York, NY: Free Press.

Bass, B. M., and Avolio, B. J. (1997). *Full Range Leadership Development: Manual for the Multifactor Leadership Questionnaire*. Palo Alto, CA: Mind Garden.

Bass, B. M., Avolio, B. J., Jung, D. I., and Berson, Y. (2003). Predicting unit performance by assessing transformational and transactional leadership. *J. Appl. Psychol.* 88, 207–218. doi: 10.1037/0021-9010.88.2.207

Bechtoldt, M. N., Welk, C., Zapf, D., and Hartig, J. (2007). Main and moderating effects of self-control, organizational justice, and emotional labour on counterproductive behaviour at work. *Eur. J. Work Organ. Psychol.* 16, 479–500. doi: 10.1080/13594320701662618

Becker, T. E., and Billings, R. S. (1993). Profiles of commitment: an empirical test. *J. Organ. Behav.* 14, 177–190. doi: 10.1002/job.4030140207

Borman, W. C., and Motowidlo, S. (1993). "Expanding the criterion domain to include elements of contextual performance," in *Personnel Selection in Organizations*, eds N. Schmitt and W. C. Borman (San Francisco, CA: Jossey-Bass).

Braun, S., Peus, C., Weisweiler, S., and Frey, D. (2013). Transformational leadership, job satisfaction, and team performance: a multilevel mediation model of trust. *Leadersh. Q.* 24, 270–283. doi: 10.1016/j.leaqua.2012.11.006

Byars, L., and Rue, R. (1994). Why performance appraisal systems fail. *Pers. Adm.* 46, 32–40.

Chang, J. W., Sy, T., and Choi, J. N. (2012). Team emotional intelligence and performance: interactive dynamics between leaders and members. *Small Group Res.* 43, 75–104. doi: 10.1371/journal.pone.0168150

Chang, T. Y. (2016). *The Relationships among Organizational Climate Organizational Trust Leadership Style Internal Control Organizational Commitment and Job Performance*. Master's thesis. Tainan: National Cheng Kung University.

Chen, K. Y., and Wang, C. H. (2011). *Advanced Statistical Analysis Using SPSS and AMOS*. Taipei: Wu-Nan Books.

Chen, S. Y. (2005). *Multivariate Analysis*, 4th Edn. Taipei: Hwa Tai Publishing Co. Ltd.

Chiu, H. J. (2006). *Quantitative Research and Statistical Analysis*. Taipei: Wu-Nan Books.

Cohen, D., and Prusak, L. (2001). *In good Company: How social Capital Makes Organizations Work*. Boston, MA: Harvard Business School Press.

Courgeau, D. (2003). *Methodology and Epistemology of Multilevel Analysis: Approaches from Different Social Sciences*. Norwell, MA: Kluwer.

Davies, M., Stankov, L., and Roberts, R. D. (1998). Emotional intelligence: in search of an elusive construct. *J. Pers. Soc. Psychol.* 75, 989–1015. doi: 10.1037//0022-3514.75.4.989

Day, A. L., and Carroll, S. A. (2004). Using an ability-based measure of emotional intelligence to predict individual performance, group performance, and group citizenship behaviours. *Pers. Individ. Differ.* 36, 1443–1458. doi: 10.1016/s0191-8869(03)00240-x

D'Errico, F. (2019). Too humble and sad: the effect of humility and emotional display when a politician talks about a moral issue. *Soc. Sci. Inf.* 58, 660–680. doi: 10.1177/0539018419893564

D'Errico, F., Bull, P., Lamponi, E., and Leone, G. (2022). Humility expression and its effects on moral suasion: an empirical study of Ocasio-Cortez's communication. *Hum. Aff.* 32, 101–117. doi: 10.1515/humaff-2022-0009

DeVellis, R. F. (2016). *Scale Development: Theory and Applications*. Newbury Park CA: Sage.

Dirks, K. T., and Ferrin, D. L. (2002). Trust in leadership: meta-analytic findings and implications for research and practice. *J. Appl. Psychol.* 87, 611–628. doi: 10.1037/0021-9010.87.4.611

Drescher, M. A., Korsgaard, M. A., Welpe, I. M., Picot, A., and Wigand, R. T. (2014). The dynamics of shared leadership: building trust and enhancing performance. *J. Appl. Psychol.* 99, 771–783. doi: 10.1037/a0036474

Druskat, V. U. (1994). Gender and leadership style: transformational and transactional leadership in the Roman Catholic Church. *Leadersh. Q.* 5, 99–119. doi: 10.1016/1048-9843(94)90023-x

Fornell, C., and Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *J. Mark. Res.* 18, 39–50. doi: 10.2307/3151312

Franco, M., and Matos, P. G. (2015). Leadership styles in SMEs: a mixed-method approach. *Int. Entrep. Manage. J.* 11, 425–451. doi: 10.1007/s11365-013-0283-2

Fu, W., and Deshpande, S. P. (2014). The impact of caring climate, job satisfaction, and organizational commitment on job performance of employees in a China's insurance company. *J. Bus. Ethics* 124, 339–349. doi: 10.1007/s10551-013-1876-y

Geringer, J. M., and Frayne, C. A. (1993). Self-efficacy, outcome expectancy and performance ofinternational joint venture general managers. *Can. J. Adm. Sci.* 10, 322–333. doi: 10.1111/j.1936-4490.1993.tb00037.x

Goh, S. K., and Low, B. Z. J. (2013). The influence of servant leadership towards organizational commitment: the mediating role of trust in leaders. *Int. J. Bus. Manage.* 9, 17–25.

Greenberg, J. (2003). Creating unfairness by mandating fair procedures: the hidden hazards of a pay-for-performance plan. *Hum. Resour. Manage. Rev.* 13, 41–57. doi: 10.1016/s1053-4822(02)00098-0

Hair, J. F., Anderson, R. E., Tatham, R. L., and Black, W. C. (1998). *Multivariate Data Analysis*, 5th Edn. Englewood Cliffs, NJ: Prentice-Hall.

Harms, P. D., and Credé, M. (2010). Emotional intelligence and transformational and transactional leadership: a meta-analysis. *J. Leadersh. Organ. Stud.* 17, 5–17. doi: 10.1177/1548051809350894

Hou, M. L. (2012). *The Study of Organizational Justice, Organizational Commitment and Trust to Willingness of Sharing Knowledge and Performance: Take TSC as an Example*. Master's Thesis. Neipu: National Pingtung University of Science and Technology.

Islam, M. N., Furuoka, F., and Idris, A. (2021). Mapping the relationship between transformational leadership, trust in leadership and employee championing behavior during organizational change. *Asia Pac. Manage. Rev.* 26, 95–102. doi: 10.1016/j.apmrv.2020.09.002

Jamal, M. (2011). Job stress, job performance and organizational commitment in a multinational company: an empirical study in two countries. *Int. J. Bus. Soc. Sci.* 2, 20–29.

Jaramillo, F., Mulki, J. P., and Marshall, G. W. (2005). A meta-analysis of the relationship between organizational commitment and salesperson job performance: 25 years of research. *J. Bus. Res.* 58, 705–714. doi: 10.1016/j.jbusres.2003.10.004

Jong, B. A., and Elfring, T. (2010). How does trust affect the performance of ongoing teams? The mediating role of reflexivity, monitoring, and effort. *Acad. Manage. J.* 53, 535–549. doi: 10.5465/amj.2010.51468649

Jung, D. I., and Avolio, B. J. (2000). Opening the black box: an experimental investigation of the mediating effects of trust and value congruence on transformational and transactional leadership. *J. Organ. Behav.* 21, 949–964. doi: 10.1002/1099-1379(200012)21:8<949::aid-job64>3.0.co;2-f

Kalsoom, Z., Khan, M. A., and Zubair, S. S. (2018). Impact of transactional leadership and transformational leadership on employee performance: a case of FMCG industry of Pakistan. *Ind. Eng. Lett.* 8, 23–30.

Kauffmann, D., and Carmi, G. (2014). How team leaders can use ICT to improve trust among virtual teams to increase collaboration. *Int. J. Eng. Innov. Technol.* 3, 204–221.

Kotsou, I., Mikolajczak, M., Heeren, A., Grégoire, J., and Leys, C. (2019). Improving emotional intelligence: a systematic review of existing work and future challenges. *Emot. Rev.* 11, 151–165. doi: 10.1177/1754073917735902

Langfred, C. W. (2004). Too much of a good thing? Negative effects of high trust and individual autonomy in self-managing teams. *Acad. Manage. J.* 47, 385–399. doi: 10.5465/20159588

Lau, C. M., and Moser, A. (2008). Behavioral effects of nonfinancial performance measures: the role of procedural fairness. *Behav. Res. Account.* 20, 55–71. doi: 10.2308/bria.2008.20.2.55

Law, K. S., Wong, C.-S., and Song, L. J. (2004). The construct and criterion validity of emotional intelligence and its potential utility for management studies. *J. Appl. Psychol.* 89, 483–496. doi: 10.1037/0021-9010.89.3.483

Lee, C. C., Lee, C. W., Jiang, L. T., Lin, C. Y., Huang, F. L., and Yu, C. (2017). The Impact of salary structure, performance requirements, and type of business on the performance of housing brokerage employees. *Asian Econ. Financ. Rev.* 7, 685–699. doi: 10.18488/journal.aefr.2017.77.685.699

Lee, C. C., and Shen, P. C. (2008). The Study of the Relationship among Organizational Structure. Job Satisfaction, Organizational Commitment and Job Performance-An Example of the Real Estate Brokers. *J. Taiwan Land Res.* 11, 65–103.

Lewis, J. D., and Weigert, A. (1985). Trust as a social reality. *Soc. Forces* 63, 967–985. doi: 10.2307/2578601

Liaw, Y. J., Chi, N. W., and Chuang, A. (2010). Examining the mechanisms linking transformational leadership, employee customer orientation, and service performance: the mediating roles of perceived supervisor and coworker support. *J. Bus. Psychol.* 25, 477–492. doi: 10.1007/s10869-009-9145-x

Liu, D., Wong, C. S., and Fu, P. P. (2012). Team leaders'emotional intelligence, personality, and empowering behavior: an investigation of their relations to team climate. *Adv. Glob. Leadersh.* 7, 77–104. doi: 10.1108/s1535-1203(2012)0000007008

MacKenzie, S. B., Podsakoff, P. M., and Rich, G. A. (2001). Transformational and transactional leadership and salesperson performance. *J. Acad. Mark. Sci.* 29:115. doi: 10.1177/03079459994506

Manaf, A. H. A., and Latif, L. A. (2014). Transformational leadership and job performance of SMEs technical personnel: the adaptability cultural approach as mediator. *Mediterr. J. Soc. Sci.* 5:648.

Masa'deh, R. E., Obeidat, B. Y., and Tarhini, A. (2016). A Jordanian empirical study of the associations among transformational leadership, transactional leadership, knowledge sharing, job performance, and firm performance: a structural equation modelling approach. *J. Manage. Dev.* 35, 681–705. doi: 10.1108/jmd-09-2015-0134

Mayer, R. C., and Gavin, M. B. (2005). Trust in management and performance: Who minds the shop while the employees watch the boss? *Acad. Manage. J.* 48, 874–888. doi: 10.5465/amj.2005.18803928

McAllister, D. J. (1995). Affect-and cognition-based trust as foundations for interpersonal cooperation in organizations. *Acad. Manage. J.* 38, 24–59. doi: 10.5465/256727

Meyer, J. P., and Allen, N. J. (1991). A three-component conceptualization of organizational commitment. *Hum. Resour. Manage. Rev.* 1, 61–89. doi: 10.1016/1053-4822(91)90011-z

Ministry of Labor (2020). *Analysis on the Real Estate Industry*. Available online at: https://www.mol.gov.tw/media/1695/al01.pdf (accessed February 20, 2020).

Mo, S., and Shi, J. (2017). Linking ethical leadership to employee burnout, workplace deviance and performance: testing the mediating roles of trust in leader and surface acting. *J. Bus. Ethics* 144, 293–303. doi: 10.1007/s10551-015-2821-z

Motowidlo, S. J., and Van Scotter, J. R. (1994). Evidence that task performance should be distinguished from contextual performance. *J. Appl. Psychol.* 79, 475–480. doi: 10.1037/0021-9010.79.4.475

Mowday, R. T., Porter, L. W., and Steers, R. (1982). *Employee—Organization Linkages: The Psychology of Commitment, Absenteeism, and Turnover*. New York, NY: Academic Press.

Nanjundeswaraswamy, T., and Swamy, D. (2014). Leadership styles. *Adv. Manage.* 7, 57–62.

Noruzy, A., Dalfard, V. M., Azhdari, B., Nazari-Shirkouhi, S., and Rezazadeh, A. (2013). Relations between transformational leadership, organizational learning, knowledge management, organizational innovation, and organizational performance: an empirical investigation of manufacturing firms. *Int. J. Adv. Manuf. Technol.* 64, 1073–1085. doi: 10.1007/s00170-012-4038-y

Nunnally, J. C., and Bernstein, I. H. (1994). *Psychometric Theory*, 3rd Edn. New York, NY: McGraw-Hill.

Podsakoff, P. M., MacKenzie, S. B., Moorman, R. H., and Fetter, R. (1990). Transformational leader behaviors and their effects on followers' trust in leader, satisfaction, and organizational citizenship behaviors. *Leadersh. Q.* 1, 107–142. doi: 10.1016/1048-9843(90)90009-7

Prati, L. M., Douglas, C., Ferris, G. R., Ammeter, A. P., and Buckley, M. R. (2003). Emotional intelligence, leadership effectiveness, and team outcomes. *Int. J. Organ. Anal.* 11, 21–40. doi: 10.1108/eb028961

Pryke, S., Lunic, D., and Badi, S. (2015). The effect of leader emotional intelligence on leader–follower chemistry: a study of construction project managers. *Constr. Manage. Econ.* 33, 603–624. doi: 10.1080/01446193.2015.1078901

Purwanto, A., Bernarto, I., Asbari, M., Wijayanti, L. M., and Hyun, C. C. (2020). Effect of transformational and transactional leadership style on public health centre performance. *J. Res. Bus. Econ. Educ.* 2, 304–314.

Rose, R. C., Kumar, N., and Pak, O. G. (2009). The effect of organizational learning on organizational commitment, job satisfaction and work performance. *J. Appl. Bus. Res.* 25, 55–66.

Schaubroeck, J. M., Peng, A. C., and Hannah, S. T. (2013). Developing trust with peers and leaders: impacts on organizational identification and performance during entry. *Acad. Manage. J.* 56, 1148–1168. doi: 10.5465/amj.2011.0358

Schlechter, A. F., and Strauss, J. J. (2008). Leader emotional intelligence, transformational leadership, trust and team commitment: testing a model within a team context. *SA J. Ind. Psychol.* 34, 42–53.

Serva, M. A., Fuller, M. A., and Mayer, R. C. (2005). The reciprocal nature of trust: a longitudinal study of interacting teams. *J. Organ. Behav.* 26, 625–648. doi: 10.1002/job.331

Setiawan, B. M., Putrawan, I. M., Murni, S., and Ghozali, I. (2016). Effect of organizational structure, leadership and trust on job performance of employee: a case study on employee at Universitas Ternama. *Int. Rev. Manage. Mark.* 6, 711–712.

Shahhosseini, M., Silong, A. D., Ismaill, I. A., and Uli, J. N. (2012). The role of emotional intelligence on job performance. *Int. J. Bus. Soc. Sci.* 3, 241–246.

Sheng, C. W., Chen, Y. E., Chou, Y. M., and Chen, M. C. (2005). A study of team structure and performance: mediated by team trust during process. *J. Hum. Resour. Manage.* 5, 75–90.

Shih, Y. S. (2006). *The Study of Transformational Leadership and Transactional Leadership on Job Performance and Turnover Intention: The Mediating Effect of Organizational Commitment*. Master's thesis. Taoyuan: National Central University.

Sholihin, M., and Pike, R. (2009). Fairness in performance evaluation and its behavioural consequences. *Account. Bus. Res.* 39, 397–413. doi: 10.1080/00014788.2009.9663374

Sosik, J. J., and Megerian, L. E. (1999). Understanding leader emotional intelligence and performance: the role of self-other agreement on transformational leadership perceptions. *Group Organ. Manage.* 24, 367–390. doi: 10.1177/1059601199243006

Staples, D. S., and Webster, J. (2008). Exploring the effects of trust, task interdependence and virtualness on knowledge sharing in teams. *Inf. Syst. J.* 18, 617–640. doi: 10.1111/j.1365-2575.2007.00244.x

Sy, T., Tram, S., and O'Hara, L. A. (2006). Relation of employee and manager emotional intelligence to job satisfaction and performance. *J. Vocat. Behav.* 68, 461–473. doi: 10.1016/j.jvb.2005.10.003

Varshney, D., and Varshney, N. K. (2017). Measuring the impact of trust on job performance and self-efficacy in a project: evidence from Saudi Arabia. *J. Appl. Bus. Res.* 33, 941–950. doi: 10.19030/jabr.v33i5.10017

Walumbwa, F. O., Avolio, B. J., and Zhu, W. (2008). How transformational leadership weaves its influence on individual job performance: the role of identification and efficacy beliefs. *Pers. Psychol.* 61, 793–825. doi: 10.1111/j.1744-6570.2008.00131.x

Wang, G., and Netemeyer, R. G. (2002). The effects of job autonomy, customer demandingness, and trait competitiveness on salesperson learning, self-efficacy, and performance. *J. Acad. Mark. Sci.* 30, 217–228. doi: 10.1177/0092070302030003003

Wen, J., Huang, S. S., and Hou, P. (2019). Emotional intelligence, emotional labor, perceived organizational support, and job satisfaction: a moderated mediation model. *Int. J. Hospital. Manage.* 81, 120–130. doi: 10.1016/j.ijhm.2019.01.009

Weng, L. C., and Li, Y. C. (2015). A multi-level investigation on the relationships between transformational leadership and knowledge sharing: the social dilemma perspective. *Int. J. Commer. Strategy* 7, 53–70.

Wong, C. S., and Law, K. S. (2002). The effects of leader and follower emotional intelligence on performance and attitude: an exploratory study. *Leadersh. Q.* 13, 243–274. doi: 10.1016/s1048-9843(02)00099-1

Yamaguchi, I. (2013). A Japan–US cross-cultural study of relationships among team autonomy, organizational social capital, job satisfaction, and organizational commitment. *Int. J. Intercult. Relat.* 37, 58–71. doi: 10.1016/j.ijintrel.2012.04.016

Yang, J., and Mossholder, K. W. (2010). Examining the effects of trust in leaders: a bases-and-foci approach. *Leadersh. Q.* 21, 50–63. doi: 10.1016/j.leaqua.2009.10.004

Yeh, H., and Hong, D. (2012). The mediating effect of organizational commitment on leadership type and job performance. *J. Hum. Resour. Adult Learn.* 8:50.

Yue, C. A., Men, L. R., and Ferguson, M. A. (2019). Bridging transformational leadership, transparent communication, and employee openness to change: the mediating role of trust. *Public Relat. Rev.* 45:101779. doi: 10.1016/j.pubrev.2019.04.012

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Lee, Li, Yeh and Yu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

APPENDIX

APPENDIX 1 | The questionnaire items.

Construct	Item (Cronbach's alpha)	Ref.
Leader emotional intelligence	Identifying the emotions of others (1) I think the manager is strongly aware of an employee's emotions. (0.726) (2) I think the manager is able to identify an employee's emotions by conversing with them. (0.696) Utilizing emotions (1) I think the manager is constantly setting up goals for employees and helping them to complete these goals. (0.719) (2) When employees encounter a problem, the manager will always give them encouragement. (0.737) Managing one's own emotions (1) I think the manager is able to control their own emotions and solve problems rationally. (0.701) (2) I think the manager has a strong ability to control their own emotions. (0.683)	Davies et al., 1998; Law et al., 2004
Transformational leadership	Idealized influence (1) The manager will tell us their core values and beliefs. (0.789) (2) I think the manager has a determined mindset. (0.812) Inspirational motivation (1) I think the manager will motivate me continuously. (0.855) (2) The manager expresses their confidence and expectations in me. (0.839) Intellectual stimulation (1) The manager encourages me to express my own opinions and ideas. (0.820) (2) The manager encourages me to ask and ponder thought-provoking questions. (0.797) Individualized consideration (1) I think the manager is considerate about the challenges I face in my work. (0.821) (2) I think the manager is willing to go out of their way to guide their employees. (0.814)	Bass, 1985; Bass and Avolio, 1997; Masa'deh et al., 2016
Transactional leadership	Contingent reward (1) When I perform well, the manager will give me positive feedback. (0.815) (2) When I achieve a better sales performance, the manager will give me extra praise. (0.751) (3) The manager will inform me of the rewards that I shall receive upon completing my tasks. (0.721) Management by exception (1) I feel that the manager delineates the shortcomings in my work. (0.739) (2) I feel that the manager pays attention to employees who fail to achieve minimum sales. (0.723)	Bass and Avolio, 1997; Masa'deh et al., 2016
Trust within a team	(1) I feel that my teammates and I can rely on each other. (0.682) (2) I feel that I trust most of my teammates. (0.669) (3) I feel that my team is impartial. (0.726) (4) I feel that my teammates will assist me when I encounter difficulties. (0.755) (5) I feel great when I accomplish goals with my teammates. (0.704)	Langfred, 2004; Staples and Webster, 2008; Jong and Elfring, 2010
Trust in supervisors	(1) I feel that the manager is impartial. (0.808) (2) I believe that the manager has excellent self-discipline. (0.791) (3) I feel that the manager would not deceive their employees for their own benefit. (0.749) (4) I feel that the manager keeps their word. (0.768) (5) I believe that the manager has good motives and intentions. (0.793)	McAllister, 1995; MacKenzie et al., 2001
Organizational commitment	Affective commitment (1) I enjoy sharing the status of my current company/branch office with other brokers. (0.506) (2) I feel that I am a part of my current company/branch office. (0.683) (3) My current company/branch office is of utmost significance to me. (0.745) Continuous commitment (1) It is not an easy task to leave my current company/branch office. (0.647) (2) If I leave the company now, my lifestyle will change drastically. (0.559) (3) I expect to stay at the current company/branch office for as long as I wish. (0.690) Normative commitment (1) I feel that the current company/branch office improves my livelihood. (0.721) (2) To me, it is immoral to switch to another company/branch office. (0.608) (3) Even if I have a better job, I would feel discontent after leaving my current company/branch office. (0.584)	Meyer and Allen, 1991; Allen and Meyer, 1996; Fu and Deshpande, 2014

(Continued)

APPENDIX 1 | Continued

Construct	Item (Cronbach's alpha)	Ref.
Job performance	Task performance (1) I feel that I am able to overcome the difficulties in my work. (0.659) (2) I take the initiative to solve problems. (0.720) (3) I do not slack even when the manager is not in the office. (0.700) Contextual performance (1) I want to be assigned or handle challenging tasks. (0.668) (2) On average, I have a considerably high work efficiency. (0.596) (3) Generally speaking, I am able to complete the tasks demanded by the company. (0.684)	Motowidlo and Van Scotter, 1994; Fu and Deshpande, 2014; Masa'deh et al., 2016

frontiers | Frontiers in Psychology

ORIGINAL RESEARCH
published: 30 May 2022
doi: 10.3389/fpsyg.2022.839064

Product-Centered Language Style in Chinese Women Online Reviews

Xian Wang[1], Congjun Mu[1], Huixian Li[1], Alison Noble[2] and Qingyi Wang[3]*

[1] College of Foreign Languages, Shanghai Maritime University, Shanghai, China, [2] Marine Engineering Department, Antwerp Maritime Academy, Antwerp, Belgium, [3] ByteDance Ltd., Beijing, China

The relationship of language style and online review has drawn increasing academic attention recently as it can provide customers with a guide to make the purchase. Extant research attaches importance to the language style that is presented in the use of function words, instead of product-related content words. This study aimed to examine the language style generated by customers' comments relating to the product based on content words, that is, product-centered language style (PCLS). We built a corpus of Chinese women clothes online reviews to explore the general picture and distinct features of PCLS and the distinct feature of PCLS. A content-word-centered Linguistic Inquiry and Word Count (LIWC) in terms of product performance is established. PCLS is calculated based on the language style matching (LSM) algorithm. Our results show that the PCLS in women clothes online review is featured by diverse and polarized language styles among three groups of women clothes buyers, and the prioritized arrangement of words of importance contributes to the PCLS. The findings benefit the women clothes industry in which it can help companies quickly find the distinctive and the transition of PCLS and offer an approach for companies to indiscriminately look into the significance of the product category from the linguistic perspective, which can help with product sale strategy and product design.

Keywords: language style, LSM, online reviews, women consumers, product content

OPEN ACCESS

Edited by:
Francesca D'Errico,
University of Bari Aldo Moro, Italy

Reviewed by:
Eduardo Moraes Sarmento,
Lusophone University of Humanities and Technologies, Portugal
Rungtai Lin,
National Taiwan University of Arts, Taiwan

****Correspondence:***
Xian Wang
Xianwang@shmtu.edu.cn

Specialty section:
This article was submitted to Language Sciences, a section of the journal Frontiers in Psychology

Received: *19 December 2021*
Accepted: *19 April 2022*
Published: *30 May 2022*

Citation:
Wang X, Mu C, Li H, Noble A and Wang Q (2022) Product-Centered Language Style in Chinese Women Online Reviews. Front. Psychol. 13:839064. doi: 10.3389/fpsyg.2022.839064

INTRODUCTION

Everyday, thousands of women consumers buy clothes based on previous online reviews (Carroll, 2008), and compared with men, their intention to purchase will increase if they receive a positive response (Garbarino and Strahilevitz, 2004). The review valence, review content, extremity and length, and language style are useful information that helps post purchaser make the decision (Forman et al., 2008; Mudambi and Schuff, 2010; Pan and Zhang, 2011). The language style of online product review has been identified as a constructive tool in management prediction, such as the third party's evaluation of participants (Romero et al., 2015), the attenuated effect of low expertise level comments (Wu et al., 2017), and the way of emotional expressions (Liu et al., 2019). However, these business applications are achieved based on the analysis of emotions inherent in the language style, but none focuses on the specific way of expression when a certain group of customers refers to the product quality. In this study, we attempted to explore the features and influences of language style relating to product quality when Chinese women clothes consumers comment online on Taobao.

The language style of online reviews has been extensively researched in recent two decades either as an information signal (Romero et al., 2015; Liu et al., 2019) or as a genre of discourse (Teso et al., 2018; Cenni and Goethals, 2021). Language style may be a tool to measure the perceived result of consumers against previous consumers (Hong et al., 2017). While Guo et al. (2020) made effort to analyze the cognitive impact based on the non-emotional content, the majority of literature relies solely on the emotional language features to describe cognitive responses to the language style of online reviews among different information receivers (Antioco and Coussement, 2018; Liu et al., 2019, Guo et al., 2020; and etc.). In addition, some scholars state that the language style of online reviews is independent of the content of reviews and put emphasis on the tokenization of emotional linguistic features, that is, the emotion-centered language style (ECLS) (Swaab et al., 2011; Ludwig et al., 2013). In the field of discourse analysis, language style is studied with a focus on the function of an acknowledged discourse instead of the internal features. For example, Ren (2018) argued that there was a mitigated language style of online reviews, and the strategy for mitigation devices is related to the positive/negative review valence. Tseng and Zhang (2020), through the example of Chinese health communication, indirectly pointed out that language style in health online reviews has its own framework.

Despite these thorough studies on the language style of online reviews (OPCs) in business research and discourse analysis, factors identified to affect the language style regarding the product are mixed in the literature. Unclassified factors such as shopping experience, price, emotions, and non-emotions (Ma and Hu, 2012) proved to contribute to language style and product-centered language style (PCLS). Guo et al. (2020) scaled down the factors by separating non-emotional from emotional content but listed a mixture of non-emotion content that contains review valence, gender, credibility, and the scientific processing system of mankind. Product attributes such as length, rating valence, and affective content directly influence the review expressions (Forman et al., 2008), which is a mixed classification as well. From these efforts, it is clear that academia intends to detail the factors inherent in the language style of online reviews that synchronize that specific way of expression, that is, the product-centered language style. However, it is not satisfactory yet as to the question: to what extent are consumers influenced by previous comments purely relating to the product? We presumed this is caused by the following: (1) it is difficult to define the product attributes, as it is related to various aspects of the product, including the performance, material, and design. Recent academic studies relating to language style in the online review are controlled by function words reflecting the emotions (Fan and Chen, 2017; Liu et al., 2018; Guo et al., 2020). (2) Compared with the description of language style from function words, which is limited in a category and in quantity, it is quite challenging to describe the product from the content words. According to Tausczik and Pennebaker (2010), content words include nouns, regular verbs, adjectives, and adverbs. (3) There are no product-centered word categories available for the comparison of language style. To make up the imbalance, this study establishes a product-centered content-word-based Linguistic Inquiry and Word Count (LIWC), a dictionary based on categories, and explores the features of the language style of women clothes, purely on product aspects (PCLS), drawing on the corpus of Chinese women clothes online reviews. In particular, we addressed the following questions:

1. Is PCLS distinguishable among women clothes buyers?
2. What linguistic features make up the PCLS variations in women online reviews?

In what follows, we first reviewed the theory about the relationship of online review writers and post readers, the studies on PCLS. In the "Methods" section, we provided an overview of the approach called language style matching (LSM) and the new LIWC based on PCLS. In the "Results" section, we demonstrated the data and relevant explanations. In the "Discussion" section, we expound on the general features of PCLS of women reviews. Finally, we concluded that the PCLS is important to monitor and diagnose the value structure of companies' products.

LITERATURE REVIEW

Language Expectancy Theory

The language used by review writers needs to consider the context of the relationship between the speaker and the audience (Pennebaker et al., 2003), as prior utterance affects the cognition of the recipients and may result in the synchronicity of language style, which is the linguistic scenario called the language expectancy theory (LET). LET posits that people develop appropriate communication styles in certain scenarios whereby they encode the language style of prior utterance, if affected, into their own communication styles (Burgoon and Miller, 1985).

The current academia in this regard focuses on comparing the language style matching between the speaker and the recipients in the context of different LET-related situations. Purposefully or unintentionally changing expectations will enhance or mitigate the effectiveness of the message whereby a given message will be incrementally enhanced, sharpening the recipients' acceptance and their way of expression (Lee and Yu, 2020). Sentence structure or word choice positively or negatively influences the expectations of recipients (Averbeck and Miller, 2014). Linguistic expectations are compared to find the proper expression for crowdfunding (Parhankangas and Renko, 2017; Koh et al., 2020). In a natural disaster, the resharing of the information on Twitter plays an incremental role in self-strengthening and self-awakening (Lee and Yu, 2020). The stylistic choices made by an individual customer are intended to provoke a particular response from other customers (Semino and Culpeper, 2011). Furthermore, words are the key aspect to show the effect of language expectations, which results in the research on semantic features. For example, language density used by patients is compared to show that higher language density means the impact of certain topics on customers is more effective (Burgoon and Miller, 1985; Burgoon et al., 2002). The word diversity relating to the language complexity can demonstrate the level of information, the higher the word diversity is, the knowledge is broader (Jensen et al., 2013;

Averbeck and Miller, 2014). In communication, people often talk as appropriately as possible to achieve their expectations (Burgoon and Miller, 1985). In terms of business online review, it is important to identify and simulate that particular way of expression to ensure conformity within that group. In general, these studies are conducted on the premise that previous utterance has contextualized the thoughts and expressions of the recipients who are faced with the same scenario; however, they do not touch on the similarities and differences of adjacent recipients that may encounter different types of a product. In this research, we used LET to explain how previous expressions relating to product quality influence the expressions of the recipients.

Product-Centered Language Style

The content of the online review is able to manifest the review value (Willemsen et al., 2011) because in e-commerce, customers rely on limited knowledge and information they have to make purchase decisions (Steenkamp and Ter Hofstede, 2002). The attributes of the content such as length, rating valence, and affective content directly influence the review expressions (Forman et al., 2008). These review content features also influence the high and low quality of reviews in the eyes of customers (Huang et al., 2015; Liu et al., 2018). Besides, online review content is a source for customers to get various knowledge about products. Through browsing the content generated by previous writers, customers have access to social knowledge about the products they are interested in (Huang and Benyoucef, 2013).

The language style inherent in the linguistic information of online reviews offers cues for customers to assess the product (Willemsen et al., 2011). The signal and observability offered by language style are one of the components to judge the quality. Linguistic variables such as word count, argument quality, review clarity, and readability are sometimes central indicators of product quality (Maslowska et al., 2017). According to Srivastava and Kalro (2019), central linguistic factors primarily consist of "'manifest' attributes such as word count, number of sentences, review rating, and number of images, which are easily observable"; "latent" factors are embedded linguistic text features, in particular language style.

Recent research still focuses on ECLS from diverse business-related factors. For example, Yin et al. (2017) found that emotional arousal from a similar language style positively affects the value of the product through the price valence. The direct language contains mental energy, feelings, and motivations and presents intense emotions in post-buyers, considering the relationship of the matching level of function words (Precht, 2000, 2003). Information and emotional process can be used to manipulate by the statement that "the language style of online reviews affects how an audience perceives the reviews, independent of the content of reviews" (Ghahtarani et al., 2020). Based on these studies, some scholars have begun focusing on the relationship between products and online reviews. Product information regarding product quality is more important than emotions, for it reduces customers' uncertainty (Pavlou and Dimoka, 2006). Useful information is obtained by customers from the non-emotional content, especially the product quality (Jiang and Benbasat, 2007; Liu et al., 2019). Customers usually add their own information to the text model to create a new comment that is in line with the previous reviews (Gernsbacher and Hargreaves, 1988; Zwaan and Radvansky, 1998). In this way, customers develop a relational tie with the cognition of the previous text in terms of the product (Wang and Karimi, 2019). The way of expression reflecting the product quality is one of the indicators that help customers make the judgment of the product quality (Guo et al., 2020). But among these studies, no one explores the features of the language style of the online review in terms of the product itself. In this study, we contributed to the literature that the PCLS is under investigation in the context of Chinese women clothes online reviews.

Studies Related to Language Style Matching

Language style matching is a technique of text analysis that measures the matching style degree of two or several people in their utterances (Ireland et al., 2011). People who have the same interest or background are likely to make utterances that mirror the sentence structure they have previously heard or read, and the verbal and non-verbal behaviors after a social interaction becomes synchronized (Ireland and Pennebaker, 2010). Such phenomenon is called LSM, which shows the matching level of identical behaviors between one individual and another, either the verbal or the written style (Gonzales et al., 2010).

Language style matching has been extensively used as an approach in various business discourse texts to study the impact and characteristics of consumer comments (Ireland et al., 2011). Using this linguistic method, it is found that figurative language of consumer reviews leads to favorable attitudes to an enjoyable purchase experience (Kronrod and Danziger, 2013); post-reviewers will note down figurative language instead of literal language to show their persuasive influence (Wu et al., 2017); the psychological processes can be identified to examine qualitative features of review text (Liu et al., 2018); mitigated perceived perception/expectancy on certain products is congruent with the negative text acquired by post-reviewers (Burgoon and Poire, 2010; Liu et al., 2018). In recent research, LSM gives rise to the function words used in an online review, by which emotions' roles are thoroughly studied (Fan and Chen, 2017; Liu et al., 2018; Guo et al., 2020). For example, the sentiment in online consumer reviews negatively or positively affects the degree of synchrony of expressed texts made by recipients (Salehan and Kim, 2016; Yin et al., 2017). The engagement in email references of the librarian's answers is measured to indicate the holistic assessment of virtual reference services (Agee, 2019). In addition to function words, the impact of topic and focus of interest on post-customers has drawn wide attention with the approach of LSM. Lin and He (2009) proposed a joint sentiment/topic model detecting how simultaneous sentiment is likely to reflect the change of topic theme, the socioeconomic characteristics. The attention focus plays a key role in families, communities, and at work when people happen to talk about a common theme (Chung and Pennebaker, 2007; Ireland et al., 2011). In this study, applying

the LSM method in women clothes online reviews in the context of PCLS, we attempted to explore the linguistic characteristics of product-centered texts, that is, what semantic expression makes one group's reviews different from those of another group.

Language style matching is an algorithm developed mainly to calculate verbal cohesiveness based on an automated textual analysis of words, which can predict the cohesiveness of the language used by various groups (Tausczik and Pennebaker, 2010). Language features including word count, sentence patterns, and verb tense were also related to the groups' cohesiveness and performance (Gonzales et al., 2010). According to Gonzales et al. (2010), the LSM algorithm is calculated in the following procedures. First, count the absolute value of the proportions of words in a given word category out of the total number of words in each group. Second, divide this value by the sum of each group, where the difference in word usage adjusted for combined word usage in that category between group 1 and group 2 is provided. Then, the result is subtracted by 1 to obtain the similarity.

In this study, we measured PCLS based on the LIWC for product content in the online reviews. Therefore, we employed a refined LSM calculation based on the formula created by Gonzales et al. (2010).

$$\text{LSM affordable} = 1 - [|(G1 - G2)|]/[(G1 + G2 + 0.000001)].$$

In the above formula, "G1" and "G2" stand for the written texts in the word category of "affordable" by each group of customers. The denominator of 0.00001 is to avoid empty sets in cases of a zero LIWC category's score. G1 is obtained by calculating the percentages of total content-and-sentiment words in the word category among the total number of words. The result is an 8-dimensional language-style score for each review.

METHODS

Overview of the Research Method

In this study, we proposed an LIWC for content words with the LSM approach in an attempt to demonstrate how the PCLS of a given group of Chinese women clothes online reviews differs from that of the other two groups. The overview of the method is seen in **Figure 1**. It is composed of two parts. The first part is to build a novel LIWC dictionary for content words, ready for the calculation of LSM. The second part is to calculate the language style of three groups of Chinese women clothes customers based on the LSM algorithm, and the results are used to compare general features of PCLS, that is, whether convergence and divergence of language style exist in the online review in terms of age group. To be specific, there are four procedures: (1) new content words representing the linguistic content words (product-centered words) are established to adapt to the online women clothes reviews; (2) the product-centered words are based on words that reflect the product categories most talked about in women online reviews; (3) the LSMs of three groups of women customer reviews are calculated; (4) the results of LSMs are compared with each other to show the level of cohesiveness and likelihood. We gave special consideration to the PCLS in terms of LSMs because customers believe that product quality can bring them a more useful assessment of the product than the sentiment does (Pavlou and Dimoka, 2006).

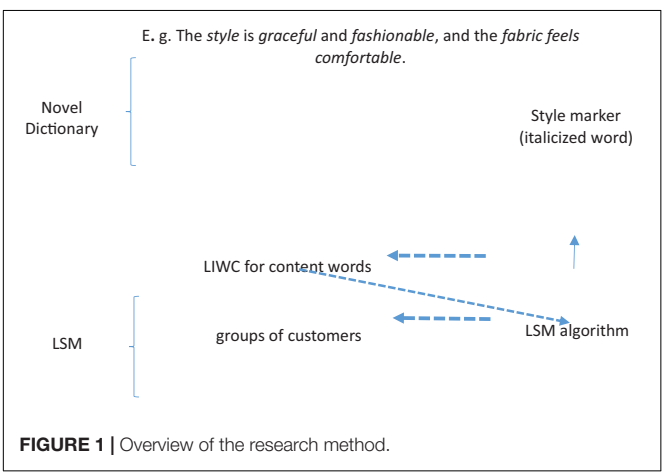


FIGURE 1 | Overview of the research method.

Reviews of Chinese Women Clothes

To make the samples as wide as possible with regards to Chinese women online reviews, we established a corpus of 32,667 online reviews written by women clothes buyers, from Taobao, an online e-commerce platform of Alibaba in China, spanning from 2018 to 2020. Information relating to product quality includes brand, manufacturer, size, color, texture, etc. More importantly, the online reviews vary with a host of factors such as age, community, and personal hobbies. Among these factors, age is one of the most influencing factors in women clothes purchases (Mirowsky and Ross, 2001). Our data are extracted based on the age group as set out by Taobao. The Taobao division of the age group is set by its own system of division, which is not in line with the division standards of the prevailing standards, for example, the standards of the World Health Organization (2010). Therefore, we followed the typical Taobao division of demography by age and classify them into three age groups: 18–25, 25–45, and 45–65, respectively, representing young, middle, and seniors. In each age group, the online reviews are expressed under considerable and enormous brands of women clothes. To address the impact of brands on online reviews, we selected those brands' reviews by sales volume not lower than the top 10 brands, which are listed on Taobao's web page. Then, the reviews are randomly taken according to brands as shown in **Table 1**.

Content-Word Linguistic Inquiry and Word Count

According to Pennebaker (2015), the LIWC is one of the core parts of LSM. LIWC is used to count and classify words. With the LIWC program, each word can be encoded and compared to see if it belongs to one of the words classified under a certain category in the dictionary. For example, "linen" is coded as a content word of material category. Then, "linen" is compared to see if "linen" is in the dictionary. LIWC produces the percentage of each LIWC category and usage frequency.

TABLE 1 | Data collection source.

Group No.	No. of reviews	Age span	Top brand(s) by volume	Age group based on Taobao division
G1	10,553	18–25	Peacebird, Tyakasha, Handu Group, La Chapelle, etc.	Young
G2	11,096	25–45	Ochirly, Prich, Hopeshow, Uniqlo, etc.	Middle
G3	11,018	45–65	Cadence, Ms. Tang, Cabenzioe, Sonmellny, etc.	Seniors

The updated LIWC program (Pennebaker, 2015) is a 90-plus category dictionary composed of almost 6,400 words, word stems, and emotions which additionally include one or more word categories or sub-dictionaries. Most prominent in the dictionary is "21 standard linguistic dimensions (e.g., percentage of words in the text that are pronouns, articles, and auxiliary verbs) and 41-word categories tapping psychological constructs (e.g., affect, cognition, biological processes, and drives)" (Pennebaker, 2015, p. 4).

The amazing part of LIWC is that it expands into aspects of words and gets updated. Psychology and children's language LIWC (Hartley et al., 2003) have been included. Recently, the causal words and insight words under the category of cognition are included as well to show the level of appraisal (Kross and Ayduk, 2008). As for the research on an online review, current LSM is based on the emotional aspect, in particular the function words of the emotional aspect in the LIWC (Leończyk, and Aleksandra., 2016; Fan and Chen, 2017). Although some scholars have argued the social influence on consumers from the perspective of economic status, gender, etc., the reliability of language style is dependent on the data from emotional functional words (Willemsen et al., 2011; Cheng and Ho, 2015). In brief, there is no LIWC categorized by words in regard to product quality at present.

According to Tausczik and Pennebaker (2010), content words and function words play an equal role in the evaluation of language style. Style or function words are those words such as pronouns, prepositions, articles, conjunctions, and auxiliary verbs. Content words include nouns, regular verbs, adjectives, and adverbs, which convey the content of communication (Haspelmath, 2001). In this study, based on the features of our data and business practice, we developed a word dictionary of content words (**Table 2**) by a combination of relevant studies (Yu, 2012; Zhang, 2019). Our dictionary is an 8-category product content dictionary, which is affordability, brand, color, clothes function, materials, size, garment style, and workmanship, respectively. For example, words such as "slim," "warm," and "young" are grouped into the category of clothes function while words such as "linen," "silk," and "texture" are classified into the material category.

TABLE 2 | Product content category used to calculate language style matching (LSM).

Dictionary category	Examples
Affordable	Cheap, worth, worthwhile
Brand	Clothes, garment
Color	Red, yellow
Clothes function	Slim, beautiful
Material	Texture, linen, silk
Size	Big, size
Clothes style	Fashion, style
Workmanship	Quality, workmanship

RESULTS

General Results

Based on our corpus of 32,667 online reviews in terms of three age groups, we acquired **Table 3**, which shows the LIWC score, LSM scores, and LSM mean.

To test the validity of these data, we conducted a standard deviation (SD) for 8-category words of three groups. The SD ranges from 0.11 to 0.5, which means that the data collected are acceptable (refer to **Figure 2**).

The Diverse and Polarized Product-Centered Language Style

The disparity of each group in terms of the specific categories of the product indicates that the PCLS of women clothes buyers is diverse in general. From **Table 3**, the LWIC within the young women (G1) shows three high figures, namely, function, style, and color, at 5.51, 4.41, and 2.57, respectively, while other aspects stay at a stunning figure lower than 1, with price at 0. This overwhelming way of expression with regards to a product indicates that young women clothes buyers, though cognitively influenced by previous customers, are likely to expressively comment on the function and style of a product than on price. Likewise, we can get the linguistic features of middle-aged women (G2) and senior women (G3). The PCLS of middle-aged women is evenly distributed with three high figures, namely, function (10.03), style (6.99), and workmanship (3.34), which manifests that middle-aged women have a more equilibrium view on clothes and tend to pose a higher demand for quality (workmanship) than young women (G1). As for G3, the PCLS is more evenly distributed with four high figures, that is, function (9.41), material (6.97), size (5.57), and style (3.14), which shows that senior women buyers hold a comprehensive look at clothes with a special concern for the size and style.

To further show the general trend of PCLS, we conducted the linear regression as shown in **Figure 3**. **Figure 3** shows the linear regression of the correlation of three LSMs. The *X*-axis stands for the three age groups, while *Y*-axis means the correlation coefficient. It is seen that the PCLS presents a downward trend in terms of three groups of women online buyers, which means that the PCLS diverges or drops in matching in terms of age groups.

TABLE 3 | Sample LSM for online review of female garments.

LIWC category	LIWC score			LSM scores			Mean LSM
	G1	G2	G3	G1/G2	G1/G3	G2/G3	
Affordable	0.00	1.82	1.39	0.00	0.00	0.87	0.29
Brand	0.74	1.22	2.79	0.75	0.42	0.61	0.59
Color	2.57	1.22	1.05	0.64	0.58	0.92	0.71
Functional	5.51	10.03	9.41	0.71	0.74	0.97	0.81
Material	0.37	2.43	6.97	0.26	0.10	0.52	0.29
Size	0.37	1.22	5.57	0.46	0.12	0.36	0.32
Style	4.41	6.99	3.14	0.77	0.83	0.62	0.74
Workmanship	0.74	3.34	1.74	0.36	0.59	0.69	0.55

The Priority of Product-Centered Language Style

In this study, we posited that the arrangement of product categories is an indicator of the language style a customer employs to demonstrate his or her identity, and in other words, the priority of texts/words relating to product categories affects the way of expression.

To acquire the PCLS of women clothes buyers in terms of priority and importance, we applied the results of LSM to the statistical analysis function of percent rank and get the importance list of the eight categories of women clothes reviews (**Table 4**).

Table 4 shows that the overall importance of the eight categories is rated based on the results of LSMs, with a breakdown of percentage from 100 to 0%. One way to interpret **Table 4** is that by the same percentage of importance, three groups show a different priority/sequence of product categories. For example, the priority of G1/G2 is clothes style, brand, clothes function, color, size, workmanship, material, and affordability, which are different from that of G1/G3, that is, clothes style, clothes function, workmanship, color, brand, size, material, and affordability. If the significance is set as 50% manually to show the high (>50%) and low attention focus (<50%), we can see that color and function are the high "categories" shared by the three groups. The second way to interpret **Table 4** is to look up the percentage of a certain product category, for example, "material." Fortunately, the category of "material" is the only one that marks the same ratio at 14.2%, signifying the "only" shared interest of focus among the three groups. As the percentage of product categories is indicative of the focus of interest from customers, we can see that the different percentage by significance shows a clear-cut priority with regards to the way of expression of women clothes reviewers.

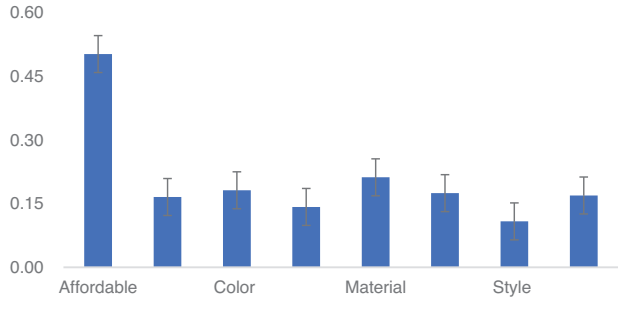


FIGURE 2 | The SD of three language style matching (LSM).

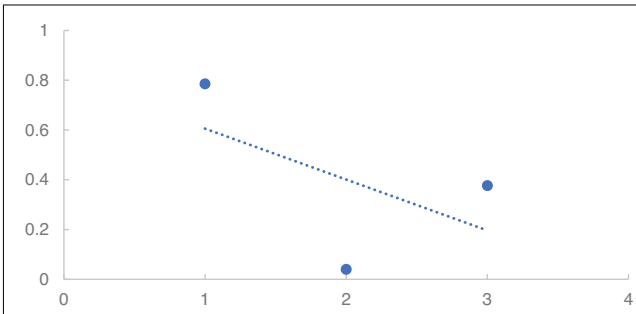


FIGURE 3 | The linear regression of correlation of three language style matching (LSM).

DISCUSSION

The first question of this study is to describe the key language features of women clothes online reviews when they comment on issues relating to a product. As shown in the previous section, diverse and polarized are the key features. We attributed the diversity of PCLS to the phenomenon called the insider language style of a speech community. Fairclough, 2003 defined insider languages as the marked language that serves to construe and reinforce the identity of a speech community. In online reviews, customers usually add their own information to the text model, which actually creates a new comment in line with the previous reviews (Gernsbacher and Hargreaves, 1988; Zwaan and Radvansky, 1998). Insider customers are sensitive to the language style of their community and follow the text rules that demonstrate such factors as the density of words and the text of social cognition (Cheng and Ho, 2015) or social variables such as gender, religion, and social position (Deckert and Vickers, 2011). To be specific, the young women clothes buyers (G1) identify themselves as the group of people who cares least about money,

TABLE 4 | Product-centered language style (PCLS) priority rankings between groups.

Rank	Significance (%)	Category	G1/G2 LSM	Category	G1/G3 LSM	Category	G2/G3 LSM
1	100.00	Clothes style	0.77	Clothes style	0.83	Clothes function	0.97
2	85.70	Brand	0.75	Clothes function	0.74	Color	0.92
3	71.40	Clothes function	0.71	Workmanship	0.59	Affordable	0.87
4	57.10	Color	0.64	Color	0.58	Workmanship	0.69
5	42.80	Size	0.46	Brand	0.42	Clothes style	0.62
6	28.50	Workmanship	0.36	Size	0.12	Brand	0.61
7	14.20	Material	0.26	Material	0.1	Material	0.52
8	0.00	Affordable	0	Affordable	0	Size	0.36

which is reinforced within the community of G1 and eventually forms a special language style.

In addition to diversity, we found that women reviews show a polarized trend by division of the group. We calculated the average mean of three LSM (0.49, 0.42, and 0.70) and then apply the correlation function. The results are 0.78, 0.04, and 0.37. By the definition of the correlation function, a higher correlation value means higher relatedness of the two variables. The G1/G3 correlation (0.04) shows that G1 and G3 have the least matching in language style, which can well explain the group disparity between young and seniors, whereas the G1/G2 correlation (0.78) shows relatively high synchrony in language style, which represents the smooth transition from young- to middle-aged woman. This result is of significance to clothes manufacturers in that products designed for young women can be branched out to middle-aged women in business promotion strategy, which may greatly increase the demography coverage of target customers.

This polarized feature can also be explained by the high and low language style matching in women clothes reviews with regards to product quality. According to Tausczik and Pennebaker (2010), the LSM score is on a scale of 0–1.0. A score of zero indicates no matching in the language style while 1.0 means high cohesiveness. Scores of 0.60 or below are usually considered low matching, and scores of 0.85 or above are deemed as signs of high matching. High and low matching levels have been proved valid in different fields. For example, high-level matching exists among customers of different nations who speak different languages when they comment on the product (Schuckert et al., 2015). Librarians can linguistically generate high-level matching feedback to the comments of the audience (Agee, 2019). The desire to enhance or mitigate the impact of language style may influence recipients who eventually produce a high or low synchronized way of expression (Giles et al., 2007). Hilte et al. (2020) proved that a certain degree of high LSM exists in two genders in teenagers' social media writing. **Table 3** shows that the only category of high matching level (>0.85) appears in G2/G3, which are clothes function, color, and price (0.97, 0.92, 0.87) while there is none in G1/G2 LSM and G1/G3 LSM. This points out that G2/G3 online reviews influence each other higher and stronger in the categories (garment function, color, and price) than other product categories, which, in the eyes of a company, may locate the "shared" aspects of a product among different customer groups.

We assumed that the declining trend of language style matching degree in **Figure 3** is largely attributed to the weakened expectancy of information or the stronger role of insider language. As language expectancy theory puts it, if people sense the style of community speech and get affected, they may translate the language style of previous comments into their own language style (Burgoon and Miller, 1985). But this LET is conditioned by the *ad hoc* fact that customers are able to identify the language style. In case of failure to the identification of prior comments, the chances of encoding are slim. Since G1 is unlikely to identify the encoded language style of G3, which may cause the low initiation of language expectancy by a host of factors such as leadership, honesty, status, age, and culture (Chung and Pennebaker, 2007), G1 may not mimic the text model of G3 and eventually comment very differently.

For the second question on the makeup of PCLS, we referred to the concept of word of significance. Groves et al. (2000) stated that the way that a potential participant responds to a product depends on the importance placed on an element of the product. It is inferred that if the importance is placed on a host of linguistic elements with different degrees of priority, and such importance is identified by the post-customer and rearranged into his online review with extra information, the importance placed on a host of elements is, in fact, the linguistic salience of prioritized product features by importance. Since people who think more about product quality will repeat the elements of product quality and be controlled in their text expression (Tausczik and Pennebaker, 2010), customers can identify and track what information they are interested in which is linked to product quality.

The fact that customers use priority PCLS to gain useful information is related to psychological linguistics. The depth of thinking or attention is reflected in the language style, which links thoughts (Tausczik and Pennebaker, 2010). In other words, the language style is reflected in thinking that correlates to how frequently a word or a group of words are mentioned. In essence, language style is closely related to attention focus, which is one of the key factors that contribute to the construction of language style (Pennebaker, 2011). The self-reference of pronouns "I" in the political advertisement is a representation of the idea of being independent (Gunsch et al., 2000). The use of "we" is linked with "better" marriage (Simmons et al., 2005). The style differences of the verb tense are the result of the speaker's attitude toward reality (Pasupathi, 2007). These studies attach importance to using tokenized words to differ the language style, but none

focuses on the arrangement of the tokenized words, let alone the PCLS priority.

It is interesting to see from **Table 4** that the ranking/sequence of G2/G3 against that of G1/G2 is 2 against 4, which represents the ratio of the language matching degree of middle-aged women/senior women vs. that of young women/middle-aged women. The higher priority of "color" in terms of middle-aged women/senior women implies that people of older ages like to dress in brighter and glossier clothes. We reasoned this to be the combination of individual focus and situational focus. The language style of a certain group is featured with the two characteristics as follows: the uniqueness of the language style of a person and the shadowed influence from other groups (Krapp, 2002; Hidi and Boscolo, 2006). Individual focus results from an individual's enthusiasm and tendency toward specific information and objects. Situational focus is activated by the surrounding environment such as the emotions of others. In the case of women of older ages, they are influenced by the contextual situation of the middle aged group (G2) in the choice of "color." Also, the fact that senior women mention more frequently the word "color" is attributed to the individual personality because the self-focus inherent in language use that "differs with age, sex, personality, and mental health" (Tausczik and Pennebaker, 2010. p. 36) can be identified from their linguistic characteristics.

CONCLUSION

Recently, scholars studied the language style used in product online reviews (Huang et al., 2015; Liu et al., 2018), solely based on the function words, which demonstrates that the emotional language style embedded in function words is related to the product. However, a large part of online reviews is concerned about product quality instead of emotions. In this study, we built a corpus of Chinese women clothes online reviews to explore the PCLS. A novel LIWC dictionary covering eight categories of product quality is sorted and established, which, in the practical sense, adds a new dimension to the existing LIWC dictionary.

With the product-category LIWC and the approach of LSM, we are able to find that the PCLS of women clothes buyers lies in (1) the diverse but polarized way of expression among different groups of buyers. The correlation function of all LSMs shows that young and seniors have the least matching in language style while middle age and seniors show a high similarity in PCLS; (2) despite the polarized language style, there exists a priority of language style relating to a certain product category, and such priority of PCLS is influenced by how customers weigh product categories in their heart. In this regard, PCLS provides a new dimension to study the language style of online reviews in the context of the product.

As pointed out, the language style of online review is to boost more effective communication. In this regard, the language style of online review may be more effective to convey customer expressions if it is used with the body signal (Poggi et al., 2013) and facial expressions (Heisler et al., 2003). This stance on effective signals may be taken into account for the future study regarding the PCLS.

This research contains practical benefits not only for women clothes industry but also for consumer management dealing with online reviews. Consumers, regardless of men or women, use language style in online consumer reviews to classify reviews and show their identity (Schindler and Bickart, 2012). The PCLS of women clothes online reviews, featuring diversity and polarity among different groups, can help companies quickly find the polarized PCLS and the transition of PCLS, which can be an aid to the product promotion strategy and product line design. Moreover, the priority PCLS offers a platform for companies to indiscriminately look into the importance of the product category from the linguistic perspective, where feedback (online reviews) of each product category can be quantitatively weighed by the same significance. In general, this quantitative PCLS can help companies identify those crucial product categories and improve product design and promotion strategy.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

XW and CM are the principal authors of this article, contributed equally to the conception of the article and to the review and critical analysis of relevant scientific literature, wrote a first version of the manuscript, arranged new sections, and produced a final version of the manuscript incorporating changes in response to the helpful suggestions of the reviewers. HL contributed to the conception, to the review and critical analysis of scientific literature on the subject, collecting data and building the corpus and she edited the figures, tables, and references in each version of manuscript. AN contributed to review the different versions of the manuscript and to polish the language. QW applied the mathematical method of percent rank function and interpreted from the perspective of social economics. All authors contributed to the article and approved the submitted version.

FUNDING

This study was supported by the National Social Science Fund of China (Grant Nos. 18CYY0231 and 17BYY103).

REFERENCES

Agee, A. (2019). Language style matching as a measure of librarian/patron engagement in email reference transactions. *J. Acad. Librar.* 45:102069. doi: 10.1016/j.acalib.2019.102069

Antioco, M., and Coussement, K. (2018). Misreading of consumer dissatisfaction in online product reviews: writing style as a cause for bias. *Int. J. Inform. Manage.* 38, 301–310. doi: 10.1016/j.ijinfomgt.2017.10.009

Averbeck, J. M., and Miller, C. (2014). Expanding language expectancy theory: the suasory effects of lexical complexity and syntactic complexity on effective

message design. *Commun. Stud.* 65, 72–95. doi: 10.1080/10510974.2013.775955

Burgoon, J. K., and Poire, B. L. (2010). Effects of communication expectancies, actual communication, and expectancy disconfirmation on evaluations of communicators and their communication behavior. *Hum. Commun. Res.* 20, 67–96. doi: 10.1111/j.1468-2958.1993.tb00316.x

Burgoon, M., and Miller, G. R. (1985). *An Expectancy Interpretation of Language and Persuasion*. London: Routledge.

Burgoon, M., Denning, V. P., and Roberts, L. (2002). Language expectancy theory. *Persuasion Handbook Dev. Theory Pract.* 2002, 117–136. doi: 10.4135/9781412976046.n7

Carroll, S. (2008). The practical politics of step-stealing and textual poaching: youtube, audio-visual media and contemporary swing dancers online. *Convergence* 14, 183–204. doi: 10.1177/1354856507087943

Cenni, I., and Goethals, P. (2021). Business responses to positive reviews online: face-work on tripadvisor. *J. Pragmat.* 180, 38–50. doi: 10.1016/j.pragma.2021.04.008

Cheng, Y. H., and Ho, H. Y. (2015). Social influence's impact on reader perceptions of online reviews. *J. Bus. Res.* 68, 883–887. doi: 10.1016/j.jbusres.2014.11.046

Chung, C., and Pennebaker, J. W. (2007). The psychological functions of function words. *Soc. Commun.* 2007, 343–359.

Deckert, S. K., and Vickers, C. H. (2011). *An Introduction to Sociolinguistics: Society and Identity*. London: A&C Black.

Fan, K. K., and Chen, Y. L. (2017). "). Exploratory study on corporate brand image and customer satisfaction on consumer purchase behavior: a case study of UNIQLO," in *Proceeding of the 2017 International Conference on Applied System Innovation (ICASI)*, (IEEE).

Fairclough, N. (2003). *Analysing Discourse: Textual Analysis for Social Research*. Hove: Psychology Press.

Forman, C., Ghose, A., and Wiesenfeld, B. (2008). Examining the relationship between reviews and sales: the role of reviewer identity disclosure in electronic markets. *Inform. Syst. Res.* 19, 291–313. doi: 10.2139/ssrn.1026893

Garbarino, E., and Strahilevitz, M. (2004). Gender differences in the perceived risk of buying online and the effects of receiving a site recommendation. *J. Bus. Res.* 57, 768–775.

Gernsbacher, M. A., and Hargreaves, D. J. (1988). Accessing sentence participants: the advantage of first mention. *J. Memory Lang.* 27, 699–717. doi: 10.1016/0749-596X(88)90016-2

Ghahtarani, A., Sheikhmohammady, M., and Rostami, M. (2020). The impact of social capital and social interaction on customers' purchase intention, considering knowledge sharing in social commerce context. *J. Innov. Know.* 5, 191–199. doi: 10.1016/j.jik.2019.08.004

Giles, H., Ogay, T., Whaley, B. B., and Samter, W. (2007). Communication accommodation in intercultural encounters. *Theor. Interc. Commun.* 2007, 293–310.

Gonzales, A. L., Hancock, J. T., and Pennebaker, J. W. (2010). Language style matching as a predictor of social dynamics in small groups. *Commun. Res.* 37, 3–19. doi: 10.1177/0093650209351468

Groves, R. M., Singer, E., and Corning, A. (2000). Leverage-salience theory of survey participation: description and an illustration. *Public Opin. Quart.* 64, 299–308. doi: 10.12688/gatesopenres.12753.1

Gunsch, M. A., Brownlow, S., Haynes, S. E., and Mabe, Z. (2000). Differential forms linguistic content of various of political advertising. *J. Broad. Electr. Med.* 44, 27–42. doi: 10.1207/s15506878jobem4401_3

Guo, J., Wang, X., and Wu, Y. (2020). Positive emotion bias: role of emotional content from online customer reviews in purchase decisions. *J. Retaili. Consumer Serv.* 52:101891. doi: 10.1016/j.jretconser.2019.101891

Hartley, J., Pennebaker, J. W., and Fox, C. (2003). Using new technology to assess the academic writing styles of male and female pairs and individuals. *J. Techn. Writ. Commun.* 33, 243–261. doi: 10.2190/9VPN-RRX9-G0UF-CJ5X

Haspelmath, M. (2001). "Word classes and parts of speech," in *International Encyclopedia of the Social and Behavioral Sciences*, eds P. B. Baltes and N. J. Smelser (Amsterdam: Elsevier).

Heisler, T., Vincent, D., and Bergeron, A. (2003). Evaluative metadiscursive comments and face-work in conversational discourse. *J. Pragm.* 35, 1613–1631.

Hidi, S., and Boscolo, P. (eds) (2006). *Writing and Motivation*. Leiden: Brill.

Hilte, L., Vandekerckhove, R., and Daelemans, W. (2020). Linguistic accommodation in teenagers' social media writing: convergence patterns in mixed-gender conversations. *J. Quant. Ling.* 2020:1807853. doi: 10.1080/09296174.2020.1807853

Hong, H., Xu, D., Wang, G. A., and Fan, W. (2017). Understanding the determinants of online review helpfulness: a meta-analytic investigation. *Dec. Support Syst.* 102, 1–11. doi: 10.1016/j.dss.2017.06.007

Huang, A. H., Chen, K., Yen, D. C., and Tran, T. P. (2015). A study of factors that contribute to online review helpfulness. *Comput. Hum. Behav.* 48, 17–27. doi: 10.1016/j.chb.2015.01.010

Huang, Z., and Benyoucef, M. (2013). From E-commerce to social commerce: a close look at design features. *Electr. Commerce Res. Appl.* 12, 246–259. doi: 10.1016/j.elerap.2012.12.003

Ireland, M. E., and Pennebaker, J. W. (2010). Language style matching in writing: synchrony in essays, correspondence, and poetry. *J. Personali. Soc. Psychol.* 99, 549–571. doi: 10.1037/a0020386

Ireland, M. E., Slatcher, R. B., Eastwick, P. W., Scissors, L. E., Finkel, E. J., and Pennebaker, J. W. (2011). Language style matching predicts relationship initiation and stability. *Psychol. Sci.* 22, 39–44. doi: 10.1177/0956797610392928

Jensen, M. L., Averbeck, J. M., Zhang, Z., and Wright, K. B. (2013). Credibility of anonymous online product reviews: a language expectancy perspective. *J. Manage. Inform. Syst.* 30, 293–324. doi: 10.2753/MIS0742-1222300109

Jiang, Z., and Benbasat, I. (2007). The effects of presentation formats and task complexity on online consumers' product understanding. *Mis Quart.* 2007, 475–500. doi: 10.2307/25148804

Koh, Y., Lee, M., Kim, J., and Yang, Y. (2020). Successful restaurant crowdfunding: the role of linguistic style. *Int. J. Contemp. Hospitali. Manage.* 32, 3051–3066. doi: 10.1108/IJCHM-02-2020-0159

Krapp, A. (2002). Structural and dynamic aspects of interest development: theoretical considerations from an ontogenetic perspective. *Learn. Instr.* 12, 383–409. doi: 10.1016/S0959-4752(01)00011-1

Kronrod, A., and Danziger, S. (2013). Wii will rock you!" the use and effect of figurative language in consumer reviews of hedonic and utilitarian consumption. *J. Consumer Res.* 40, 726–739. doi: 10.1086/671998

Kross, E., and Ayduk, O. (2008). Facilitating adaptive emotional analysis: distinguishing distanced-analysis of depressive experiences from immersed-analysis and distraction. *Personali. Soc. Psychol. Bull.* 34, 924–938. doi: 10.1177/0146167208315938

Lee, C. H., and Yu, H. (2020). the impact of language on retweeting during acute natural disasters: uncertainty reduction and language expectancy perspectives. *Industr. Manage. Data Syst.* 120, 1501–1519. doi: 10.1108/IMDS-12-2019-0711

Leończyk, and Aleksandra. (2016). The role of language proficiency in the perception of l2 taboo words by late bilingual speakers. *Eur. Yearbook* 16, 25–58. doi: 10.1075/eurosla.16.02leo

Lin, C., and He, Y. (2009). "Joint sentiment/topic model for sentiment analysis," in *Proceeding of The 18th ACM Conference on Information and Knowledge Management 09*, 375.

Liu, A. X., Xie, Y., and Zhang, J. (2019). It's not just what you say, but how you say it: the effect of language style matching on perceived quality of consumer reviews. *J. Interact. Mark.* 46, 70–86. doi: 10.1016/j.intmar.2018.11.001

Liu, S. Q., Bogicevic, V., and Mattila, A. S. (2018). Circular versus angular service scape: "shaping" customer response to a fast service encounter pace. *J. Bus. Res.* 89, 47–56. doi: 10.1016/j.jbusres.2018.04.007

Ludwig, S., De Ruyter, K., Friedman, M., Brüggen, E. C., Wetzels, M., and Pfann, G. (2013). More than words: the influence of affective content and linguistic style matches in online reviews on conversion rates. *J. Mark.* 77, 87–103. doi: 10.1509/jm.11.0560

Ma, Y. L., and Hu, Z. M. (2012). Online review of research in 2006-2011: thematic analysis and trend forecast. *Econ. Manage. Rev.* 28, 69–75.

Maslowska, E., Malthouse, E. C., and Viswanathan, V. (2017). Do customer reviews drive purchase decisions? The moderating roles of review exposure and price. *Dec. Support Syst.* 98, 1–9. doi: 10.1016/j.dss.2017.03.010

Mirowsky, J., and Ross, C. E. (2001). Age and the effect of economic hardship on depression. *J. Health Soc. Behav.* 2001, 132–150. doi: 10.2307/3090174

Mudambi, S. M., and Schuff, D. (2010). Research note: what makes a helpful online review? A study of customer reviews on Amazon. com. *MIS quarterly* 2010, 185–200. doi: 10.0000/

Pan, Y., and Zhang, J. Q. (2011). Born unequal: a study of the helpfulness of user-generated product reviews. *J. Retaili.* 87, 598–612. doi: 10.1016/j.jretai.2011.05.002

Parhankangas, A., and Renko, M. (2017). Linguistic style and crowdfunding success among social and commercial entrepreneurs. *J. Bus. Vent.* 32, 215–236. doi: 10.1016/j.jbusvent.2016.11.001

Pasupathi, M. (2007). Telling and the remembered self: linguistic differences in memories for previously disclosed and previously undisclosed events. *Memory* 15, 258–270. doi: 10.1080/09658210701256456

Pavlou, P. A., and Dimoka, A. (2006). The nature and role of feedback text comments in online marketplaces: implications for trust building, price premiums, and seller differentiation. *Inform. Syst. Res.* 17, 392–414.

Pennebaker, J. W. (2011). *The Secret Life of Pronouns: What Our Words Say About Us*. New York: Bloomberry Press.

Pennebaker, J. W. (2015). *The Development and Psychometric Properties of LIWC 2015*. Austin: University of Austin.

Pennebaker, J. W., Mehl, M. R., and Niederhoffer, K. G. (2003). Psychological aspects of natural language use: our words, our selves. *Ann. Rev. Psychol.* 54, 547–577. doi: 10.1146/annurev.psych.54.101601

Poggi, I., D'Errico, F., and Vincze, L. (2013). Comments by words, face and body. *J. Multi. User Int.* 7, 67–78.

Precht, K. (2000). *Patterns of Stance in English*. Flagstaff, AZ: Northern Arizona University.

Precht, K. (2003). Stance moods in spoken english: evidentiality and affect in british and american conversation. *Text Talk* 23, 239–257. doi: 10.1515/text.2003.010

Ren, W. (2018). Mitigation in chinese online consumer reviews. *Dis. Context Med.* 26, 5–12. doi: 10.1016/j.dcm.2018.01.001

Romero, D. M., Swaab, R. I., Uzzi, B., and Galinsky, A. D. (2015). Mimicry is presidential: linguistic style matching in presidential debates and improved polling numbers. *Personali. Soc. Psychol. Bull.* 41, 1311–1319. doi: 10.1177/0146167215591168

Salehan, M., and Kim, D. J. (2016). Predicting the performance of online consumer reviews: a sentiment mining approach to big data analytics. *Dec. Support Syst.* 81, 30–40. doi: 10.1016/j.dss.2015.10.006

Semino, E., and Culpeper, J. (2011). "Stylistics," in *Pragmatics in Practice. Handbooks of Pragmatics*, eds J.-O. Östman and J. Verschueren (Amsterdam: John Benjamins).

Schindler, R. M., and Bickart, B. (2012). Perceived helpfulness of online consumer reviews: the role of message content and style. *J. Consumer Behav.* 11, 234–243. doi: 10.1002/cb.1372

Schuckert, M., Liu, X., and Law, R. (2015). Hospitality and tourism online reviews: recent trends and future directions. *J. Travel Tourism Mark.* 32, 608–621. doi: 10.1080/10548408.2014.933154

Simmons, R. A., Gordon, P. C., and Chambless, D. L. (2005). Pronouns in marital interaction: what do "you" and "i" say about marital health? *Psychol. Sci.* 16, 932–936. doi: 10.1111/j.1467-9280.2005.01639.x

Srivastava, V., and Kalro, A. D. (2019). Enhancing the helpfulness of online consumer reviews: the role of latent (content) factors. *J. Int. Mark.* 48, 33–50. doi: 10.1016/j.intmar.2018.12.003

Steenkamp, J. B. E., and Ter Hofstede, F. (2002). International market segmentation: issues and perspectives. *Int. J. Res. Mark.* 19, 185–213. doi: 10.1016/S0167-8116(02)00076-9

Swaab, R. I., Maddux, W. W., and Sinaceur, M. (2011). Early words that work: when and how virtual linguistic mimicry facilitates negotiation outcomes. *J. Exp. Soc. Psychol.* 47, 616–621. doi: 10.1016/j.jesp.2011.01.005

Tausczik, Y. R., and Pennebaker, J. W. (2010). The psychological meaning of words: LIWC and computerized text analysis methods. *J. Lang. Soc. Psychol.* 29, 24–54. doi: 10.1177/0261927X09351676

Teso, E., Olmedilla, M., Martínez-Torres, M. R., and Toral, S. L. (2018). Application of text mining techniques to the analysis of discourse in eWOM communications from a gender perspective. *Technol. Forecast. Soc. Change* 129, 131–142. doi: 10.1016/j.techfore.2017.12.018

Tseng, M. Y., and Zhang, G. (2020). Perceptions of and attitudes toward elastic language in online health communication in chinese. *Lingua* 233:102750. doi: 10.1016/j.lingua.2019.102750

Wang, F., and Karimi, S. (2019). This product works well (for me): the impact of first-person singular pronouns on online review helpfulness. *J. Bus. Res.* 104, 283–294. doi: 10.1016/j.jbusres.2019.07.028

Willemsen, L. M., Neijens, P. C., Bronner, F., and De Ridder, J. A. (2011). Highly recommended!" the content characteristics and perceived usefulness of online consumer reviews. *J. Comput. Med. Commun.* 17, 19–38. doi: 10.1111/j.1083-6101.2011.01551.x

World Health Organization (2010). *Global Recommendations on Physical Activity for Health*. Geneva: World Health Organization.

Wu, L., Shen, H., Fan, A., and Mattila, A. S. (2017). The impact of language style on consumers' reactions to online reviews. *Tourism Manage.* 59, 590–596. doi: 10.1016/j.tourman.2016.09.006

Yin, D., Bond, S. D., and Zhang, H. (2017). Keep your cool or let it out: nonlinear effects of expressed arousal on perceptions of consumer reviews. *J. Mark. Res.* 54, 447–463. doi: 10.1509/jmr.13.0379

Yu, G. R. (2012). Questionnaire preparation and dimension construction of clothing brand evaluation. *Silk* 4, 65–69.

Zhang, Y. Z. (2019). The weight calculation of garment aesthetic evaluation dimension based on analytic hierarchy process. *Wool Textile Technol.* 47, 81–85.

Zwaan, R. A., and Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychol. Bull.* 123:162. doi: 10.1037/0033-2909.123.2.162

Conflict of Interest: QW was employed by ByteDance Ltd.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Wang, Mu, Li, Noble and Wang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers | Frontiers in Communication

ORIGINAL RESEARCH
published: 09 June 2022
doi: 10.3389/fcomm.2022.869084

Effectiveness, Attractiveness, and Emotional Response to Voice Pitch and Hand Gestures in Public Speaking

Emma Rodero [1,2]*

[1] Media Psychology Lab, Department of Communication, Pompeu Fabra University, Barcelona, Spain, [2] UPF-Barcelona School of Management, Barcelona, Spain

Some non-verbal cues, such as voice pitch and gestures, can influence the individual's emotional response to different messages. For this reason, non-verbal communication can be understood as the language of emotion, while the content is the language of ideas. But the research question here is what voice pitch and hand gesture strategies are more effective, attractive, and elicit a stronger emotional response. To answer this question, this study analyzes some pitch and hand gesture strategies in public discourses. In the experiment, 120 participants were exposed to different public presentations with three conditions for voice pitch (few, medium, and many variations) and three for hand gestures (smooth, medium, and strong intensity). Then, they rated the effectiveness and attractiveness levels and self-report arousal and valence while their emotional response was registered. Physiological arousal was measured with electrodermal activity (EDA) and valence with facial expression recognition. Participants perceived the medium pitch variations and the medium intensity of gestures as the most effective and attractive strategies with the highest self-reported arousal and valence. The combination that elicited the strongest autonomic arousal was medium pitch variations-strong gestures. Finally, the most positive emotional valence was for many pitch variations and medium gestures and the happiest for the combination many-strong.

Keywords: voice pitch, hand gestures, effectiveness, attractiveness, emotional response, arousal, valence

OPEN ACCESS

Edited by:
Oliver Niebuhr, University of Southern Denmark, Denmark

Reviewed by:
Melisa Stevanovic, Tampere University, Finland
Kerstin Fischer, University of Southern Denmark, Denmark

***Correspondence:**
Emma Rodero
emma.rodero@upf.edu

Specialty section:
This article was submitted to Language Sciences, a section of the journal Frontiers in Communication

Received: *03 February 2022*
Accepted: *12 April 2022*
Published: *09 June 2022*

Citation:
Rodero E (2022) Effectiveness, Attractiveness, and Emotional Response to Voice Pitch and Hand Gestures in Public Speaking. Front. Commun. 7:869084. doi: 10.3389/fcomm.2022.869084

INTRODUCTION

When we see and listen to a person talking, two different parts can be distinguished in communication: what this person says and how they say it, or in other words, the content and the form of the message. How this person says something or the form of the message, is called non-verbal communication, and it is as important or more as speech content (Birdwhistell, 1970; Guyer et al., 2019). Every part of our body, each movement, facial expression, or pitch variation have a meaning (Burgoon et al., 2010). In fact, our brain can draw impressions about a speaker in milliseconds, just by looking at the speakers' faces and bodies (Bar et al., 2006) or by listening to their voices.

Knapp (1980) provided a classification of seven dimensions of non-verbal communication. These components are kinesics (gestures, movement, and postures), paralanguage (related to voice features), proxemics (space and distances), physical contact (touching), physical characteristics of people (skin color…), adornments (clothes, jewelry…), and environmental factors (related to the physical setting). Of all of them, the most relevant codes are kinesics and vocalics (Burgoon et al., 2010). Therefore, in this study, we will focus on these two categories.

There are some studies about non-verbal communication analyzing the different components, especially kinesics. The most productive research period was between 1970 and 1980. However, according to Gordon et al. (2006), much of this research has a theoretical research orientation or an application demonstration orientation, with no empirical data to support the analysis and conclusions. Along with a lack of empirical research, few studies have employed psychophysiological methods to know the underlining mechanisms of these codes and how they are cognitively processed. Using these techniques applied to the study of non-verbal communication has some advantages (Potter and Bolls, 2012; Rodero, 2021). The first one is that the response is registered at the same time that people are exposed to the non-verbal stimulus. Therefore, we do not have to trust in individuals' perceptions, in what they think or perceive when they are not in front of the stimulus. The second is that we can measure the individuals' unconscious or implicit responses. This benefit is very important when analyzing non-verbal communication, as most signals are processed unconsciously. The third is that we can have a more objective measure of the non-verbal stimulus than trusting self-reporting. Consequently, it is essential to include these techniques to examine non-verbal signals, add rigor to the analysis, and have a better idea of what subjects are thinking and feeling.

Non-verbal communication cues can influence the message's perception and processing (Newman et al., 2016; King et al., 2020). We use our body and voice changes to reinforce or qualify what we are saying, convey emotions, attitudes, and intentions, regulate the flow of communication, establish contact, and provide feedback with/to other people. Clark and Greatbatch (2011) have shown that charismatic leaders use voice pitch variations, eye contact, gestures, and facial expressions. Consequently, how presenters in public speaking use their voices and gestures is crucial to engage the audience, attract attention, and elicit different emotions (Jackob et al., 2011; Talley and Temple, 2015). Scott (1920), in an old document titled Elocution, defined nonverbal communication as the language of sentiment and emotion, whereas the content was the language of ideas. As voice pitch and gestures can be relevant in determining how effective and attractive a public speech is (Beattie and Shovelton, 2005; Dargue et al., 2019) and elicit emotional reactions, we will focus on these features and variables in this study, analyzed with self-report and psychophysiological techniques. Therefore, this research aims to analyze the effectiveness, attractiveness, and emotional response (arousal and valence) of different strategies, such as pitch variations and intensity of hand gestures, in public discourses.

The main contributions of this study will be 2-fold: a) to advance in the analysis of non-verbal cues, such as voice pitch and gestures and their influence on information processing and emotional response; b) to establish some practical recommendations to use these two features when delivering public speeches.

Voice Pitch

Our voice plays a huge role in our social relationships (Cherry, 1953; Belin et al., 2004) and, therefore, in persuasive messages as public discourses (Rodero et al., 2022). How we use it is called prosody. Prosody is very important in non-verbal communication as it represents the set of features that we use when speaking. If the voice is our instrument, prosody metaphorically could be the song we play. Prosody is a linguist component with an essential role in communication, as it can influence the perception of a speaker (Elbert and Dijkstra, 2014; Varghese and Nilsen, 2020; Weinstein et al., 2020) and affect the cognitive processing of messages (Hirschberg and Pierrehumbert, 1986; Levi and Pisoni, 2007; Rodero et al., 2017). How a person speaks can be very relevant for the level of effectiveness and the listeners' cognitive and emotional response.

Prosody is composed of intonation, stress, and rhythm, among other features (Wells, 2007; Rodero et al., 2019). One of the main components is intonation. Intonation is composed of the different pitch variations (rises and falls) produced in spoken language (Tench, 2015). However, research about the influence of these tone or pitch variations in public speaking is scarce despite its importance. The studies about their impact on persuasive messages indicate that the most effective strategy is using a moderate pitch level with significant pitch variations, especially against monotonous styles (Burgoon et al., 1990; Hincks, 2004; Strangert, 2005; Yang et al., 2020). Jackob et al. (2011) found that substantial pitch changes increased persuasiveness and credibility. Rosenberg and Hirschberg (2009), Yang et al. (2020), and Niebuhr and Silber-Varod (2021) concluded that a higher and varied pitch is related to charisma. Rodero et al. (2017) demonstrated that the strategies with pitch variations in commercials increased listeners' attention and the ads were better recalled. Moreover, the messages with changes were perceived as the most effective and adequate. Recently, in public discourses, Rodero et al. (2022) showed that a moderate emphatic intonation was considered the most effective and credible, grabbed more attention, and elicited higher autonomic arousal. Conversely, other authors have not found pitch effects on persuasion (Elbert and Dijkstra, 2014), credibility (Chebat et al., 2007), or the attitude toward the message (Gélinas-Chebat et al., 1996).

The reason why pitch variations can work better from a cognitive point of view is that tone changes can more easily grab the listeners' attention and improve comprehension compared to a lack of variations or less variety, according to the Principle of Distinctive and Contrastive Coherence of Prosody (Rodero, 2015). When the spoken sequence is produced with pitch changes, there is an acoustic contrast between intonation rises and falls. Then, this sequence sounds more dynamic (Addington, 1968) and, hence, more effective (Rodero et al., 2017). But

what is more important is that this auditory difference may attract the listener's attention and increase arousal, especially in high tones or rising contours (Hill and Miller, 2010). Together with this contrastive effect, these changes also act by differentiating the relevance of the content (less important in high pitch and more important in low pitch) and, consequently, beneficiating comprehension. However, although these changes are essential, an excessive number of variations could be counter-productive, as the sequence could sound exaggerated or unnatural. Therefore, theoretically, a moderate strategy with progressive pitch variations could be the best option (Rodero et al., 2022). On the contrary, if there are not many pitch changes, the sequence can sound monotonous and dull; thus, less effective and attractive (Glass, 1991; Knapp and Hall, 2007; Rodero, 2013). An unsuccessful speaker is monotonous (Strangert and Gustafson, 2008). Moreover, this lack of variation may affect arousal and attention due to the perception of a constant and repetitive sound. It is the phenomenon called "sensory adaptation" (Wark et al., 2007). Our brain becomes fatigued, as there is no acoustic contrast. Therefore, this monotony could hinder the emotional response, reducing participants' arousal, and being perceived as more negative. Based on these studies and ideas, we formulate the two first hypotheses of this research.

H1: The style with medium pitch variations will be perceived as more effective and attractive than few and many changes.
H2: The style with medium pitch variations will elicit higher arousal and positive valence than few and many changes, both in self-report and physiological and facial expression analysis.

Hand Gestures

With prosody, gestures are important for a message's perception and information processing (King et al., 2020; Rodero et al., 2022). Inside this category, hand gestures are the spontaneous movements individuals make with their hands while talking (Wagner et al., 2014). These movements usually accompany speech (Iverson and Goldin-Meadow, 1998). According to Ekman and Friesen (1969), there are five kinds of gestures: illustrators (that convey semantic content), emblems (that represent conventional signs), regulators (that control conversational flow), emotional displays (for expressing emotions), and adaptors (to contact and touch). Most research has been devoted to analyzing illustrators (Harris, 2003). These gestures are coordinated with speech and aligned with semantic content (Jannedy and Mendoza-Denton, 2005). As this research also explores voice pitch, a feature related to speech, illustrators will be the gestures examined here.

Hand gestures, especially illustrators and regulators, aid lexical and semantic processing (Krauss et al., 1991) and provide some information to the listener (Kendon, 1981). Therefore, its use has shown to be beneficial for perception and cognitive processing, especially compared to the lack of them (Ekman et al., 1980; Loehr, 2004). For perception, people that use hand gestures are considered as more effective (Beattie and Shovelton, 2005), persuasive (Jackob et al., 2011; Peters and Hoetjes, 2017), credible (Maricchiolo et al., 2009), dominant (Gnisci and Pace, 2014), extrovert (Neff et al., 2010), sociable (Burgoon et al., 1990), and honest (King et al., 2020). Therefore, there is a positive perception when someone uses hand movements. For cognitive processing, some studies have shown that these gestures can improve understanding and comprehension (Graham and Argyle, 1975; Riseborough, 1981; Hostetter, 2011; Kelly et al., 2015; Dargue et al., 2019), attention (Berger and Popelka, 1971), memory (Ianì and Bucciarelli, 2017; Clough and Duff, 2020), stimulate emotional reactions (Jackob et al., 2011), and clarify the discourse content (Holle and Gunter, 2007). According to the integrated system hypothesis (Kelly et al., 2010), gestures are part of an integrated system together with speech. They illustrate and draw what we are saying; therefore, they can help to understand the message.

Most research has compared the use against not use of hand gestures. For instance, Maricchiolo et al. (2009) showed that effectiveness was higher in the gesture presence condition than in the no-gesture condition. However, studies about the way to use them are scarce. Specifically, the number or intensity of gestures can be the determinant variables that affect perception and emotional response, as they can modify the expressivity level. In this regard, some studies have demonstrated positive results for a high number of gestures. Gnisci and Pace (2014) found that some speakers were considered more dominant when making a high rate of gestures. Along with this, a study of the well-known TED talks concluded that hand gestures made the speaker seem more compelling. In the most popular talks, the speakers used 465 hand gestures, whereas in the least popular, the speakers used 272 gestures (Van Edwards, 2017). However, these were self-report and perception studies that did not measure the emotional response and did not use psychophysiological techniques. The most similar study to this research was conducted recently by Rodero et al. (2022). These authors compared three different amounts of gestures. They concluded that using a moderate quantity of gestures was perceived as the most effective and credible, and a high quantity of gestures increased the participants' attention and body activation. But, along with the number of gestures, the intensity of these hand movements could also be an influent factor, as this study analyzes. A person moving both hands simultaneously could be considered more expressive than an individual moving gently only one hand when speaking. However, as with voice pitch, an excessive intensity in making gestures might be exaggerated and provoke distractions, especially as these features are visually registered. Therefore, this research hypothesizes that the medium intensity of gestures (not smooth but not strong) could be the best strategy. Strongly marked gestures could distract the audience, whereas smooth gestures could not elicit a significant emotional response. In consequence, we posit the last hypotheses of this study.

H3: The style with medium intensity of gestures will be perceived as more effective and attractive than the styles with smooth and strong gestures.
H4: The style with medium intensity of gestures will elicit higher arousal and positive valence than the styles with smooth and

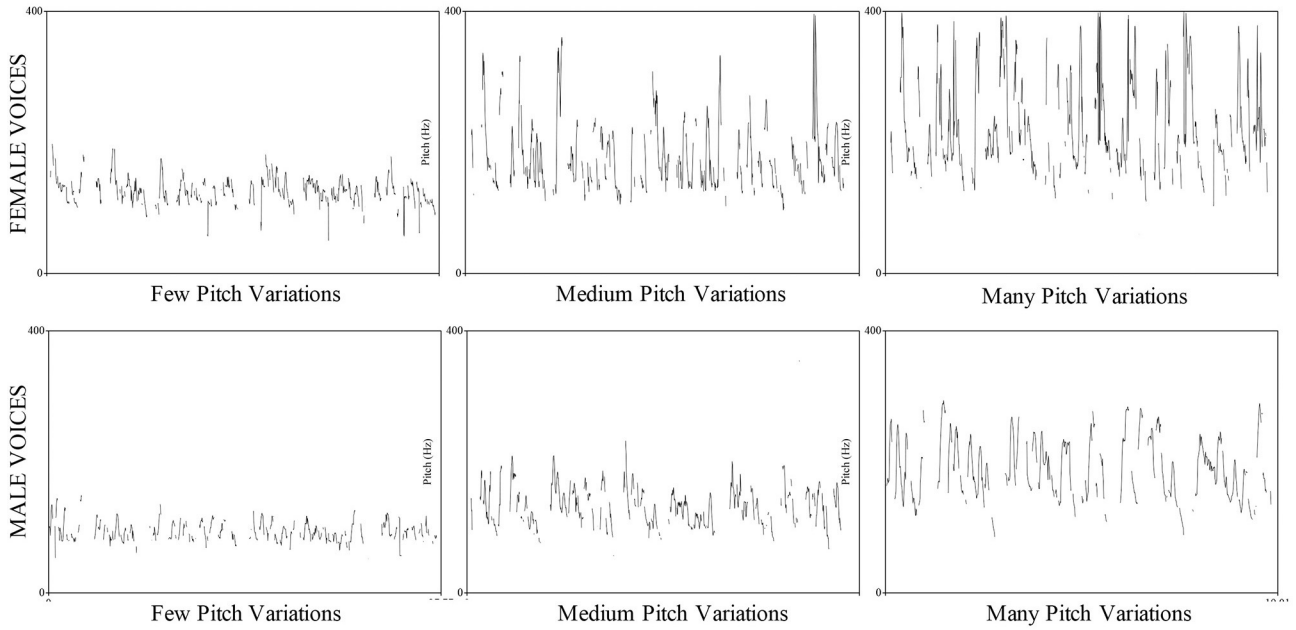


FIGURE 1 | Intonation.

strong gestures, both in self-report and physiological and facial expression analysis.

MATERIALS AND METHODS

The design of this study was a 3-voice pitch variation (few–medium–many) by 2 (scripts) within-subject experiment applied to effectiveness and self-reported valence and arousal. The design was the same for gestures but with three different intensities (smooth–medium–strong). The statistical analyses of physiological data (arousal) and valence were submitted to a 3-pitch variation (voice pitch styles or gestures) ANOVA. The voice pitch and gestures styles and actors were within-subject factors, whereas the two discourses were between-subject factors.

Stimuli

A total of forty-eight videos were recorded for this experiment, 24 on each speech, 12 by each of the four speakers. The discourses were prepared specifically for this study to control all the variables. The content was neutral and familiar for college students. The discourses were short as elevator pitches with an average of 25 s. Actors recorded both speeches with a different combination of non-verbal features: three pitch variations (few–medium–many) and three intensities of gestures (smooth–medium–strong).

Voice pitch was analyzed using Praat's acoustic analysis software (Boersma and Weenink, 2022) by measuring pitch level and pitch range. There were three different types of pitch variations (few–medium–many). The pitch values were calculated in semitones (re 1Hz). The pitch level average in the strategy with few pitch variations was 80.68 st. In the strategy with medium variations, the pitch level average in actors was 85.23 st. Finally, in many variations, the level was 90.82 st. There were significant differences among the three pitch variations (F (2,9) $= 13.98$, $p = 0.002$). The pitch range or the variation between the maximum and the minimum pitch was 19.85 st. in the strategy with few pitch variations, 24.13 st. in medium variations, and 29.33 st. in the many variations style. There were significant differences among the three pitch variations (F (2,9) $= 6.85$, $p = 0.016$). **Figure 1** shows the curves of intonation with the same text.

The intensity of gestures was controlled and measured with the software Codimg, a performance analysis application that classifies gestures. The gestures used in this study were illustrators (Ekman and Friesen, 1969). The movement was slight in the smooth strategy with one hand in the low part of the body, more intense with one hand in the high part of the body for the medium strategy, and more marked using two hands chest-high for strong gestures, as shown in **Figure 2**. Accordingly, facial expressions were growing in intensity from smooth to strong.

The discourses were recorded in a TV studio equipped with a teleprompter for the actors to read the speeches, similar lights and acoustic conditions, a professional lavalier microphone, and a blue background. In postproduction, all the audios were normalized by using Pro Tools.

Participants

A total of one hundred and twenty university students (79 women and 41 men) aged between 20 and 30 years ($M = 21$, $SD = 1.08$) formed the sample of this study. Participants were communication students. They were divided into two groups. A total of sixty watched the first discourse (24 videos) and 60 the

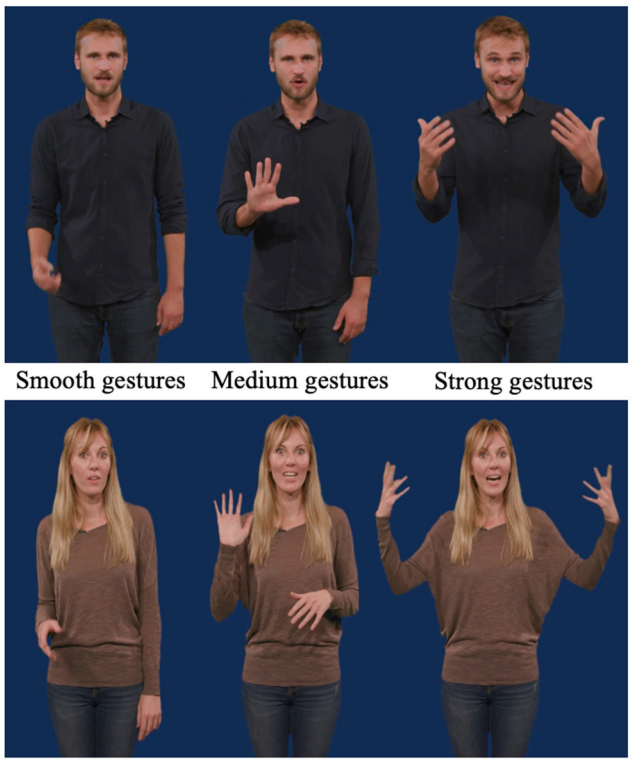


FIGURE 2 | Gestures.

second one (24 videos). The 120 participants were exposed to all the experimental conditions except the type of discourse: 3 voice pitch variations (few–medium–many) and 3 intensity gestures (smooth–medium–strong) by 4 actors.

Dependent Variables

For effectiveness, we partially used a scale based on the Rodero et al. (2013) effectiveness index. The scale measured how natural, clear, and understandable the speech was. These three dimensions were measured on a 5-point scale, with 1 representing the minimum value and five the maximum. The scale had a high Cronbach's alpha coefficient ($\alpha = 0.83$).

Participants rated how competent, persuasive, and pleasant the speech was for attractiveness. These three dimensions were measured on a 5-point scale, with 1 representing the minimum value and five the maximum. The scale had a high Cronbach's alpha coefficient ($\alpha = 0.90$).

The emotional response was analyzed using their two main dimensions: arousal and valence (Bradley and Lang, 1994), both in self-report and physiological (arousal) and facial recognition (valence) response. The first dimension, arousal, is responsible for body activation and is the state of being awake and alert (Dawson et al., 2000). This state is produced as sympathetic nervous system response. If the level of arousal is high, there is a strong emotional activation. The second dimension, valence, determines whether this emotion is pleasant or attractive (positive) or unpleasant or avert (negative). When individuals are exposed to a stimulus, they can react positively (approaching) by activating the appetitive system or, on the contrary, with rejection by activating the aversive system (Lang and Bradley, 2013). For example, happiness is a positive emotion, while sadness is a negative one.

To measure arousal and valence in self-report, the Self-Assessment Manikin (SAM) scale was used (Bradley and Lang, 1994). Self-report arousal, defined as the intensity of the emotion, was a 5-point scale. Self-report valence, how positive and negative the emotion is, was registered on two 5-point scales: one for positive and one for negative. The final valence was computed by subtracting the positivity score from the negativity score.

Physiological arousal was measured in this study by registering the electrodermal activity. This technique measures the changes in the skin's electrical properties in response to the activity of the eccrine glands. If individuals are developing a complex task or being nervous, the sympathetic nervous system activates the body, increasing sweat secretion. The electrodermal activity registers the changes in the skin conductance provoked by sweating by applying a constant voltage. If the individuals are aroused, they are better electrical conductors, and the response is higher. Therefore, electrodermal activity is considered a measure of the sympathetic nervous system's activity to register the motivational response and increase emotional activation or arousal (Potter and Bolls, 2012). In this study, the electrodermal activity (EDA) was quantified by measuring the conductance of a constant voltage (0.5V) delivered to two 8-mm AG/AGCL electrodes placed on the participants' fingers of the non-dominant hand and connected to a Biopac bioamplifier MP-160. The signal was recorded with a sampling rate of 1,000 Hz and low pass filters.

Apart from self-report, valence can be measured by different means, such as electromyography or software for facial gesture recognition. The emotional valence is analyzed in this study using Facereader, a software for facial gesture recognition. This software measures facial movements to identify emotional expressions based on Ekman's face studies. The analysis indicates the emotional valence (positive or negative), and the kind of emotion, among other parameters.

Procedure

Participants arrived at the laboratory, and they read and signed the consent form. Then, the electrodermal activity sensors were attached to the participant's fingers, and they were sitting in a comfortable armchair. Participants used headphones to listen to the videos. The order was random. After watching each video, participants had to rate the effectiveness and attractiveness levels. This procedure was repeated with all the videos. When participants finished, the sensors were removed, and they were dismissed. The experiment was conducted in two different sessions of 40 min (speech one and speech two). Students received two credits for their participation.

RESULTS

Effectiveness and Attractiveness

To test H1 and H3, a 3-pitch variation (few–medium–many) by 3 intensities of gestures (smooth–medium–strong) by 2 scripts (discourse one and discourse two) factorial MANOVA was performed on the effectiveness and attractiveness variables. There were no significant differences in scripts in any of them. The combination resulted in significant main effects for voice pitch in effectiveness, $F_{(2, 118)} = 1117$, $p < 0.001$, partial $\eta^2 = 0.139$. The medium style was the most effective, followed by many and few variations. The *post hoc* tests (Tukey) showed significant differences among all the conditions. Also, there were significant differences for attractiveness, $F_{(2, 118)} = 1682$, $p < 0.001$, partial $\eta^2 = 0.195$. The most attractive style was the strategy with medium variations, followed by many and few changes. The significant differences were among all the conditions, according to the *post hoc* tests.

Regarding gestures, the combined dependent variables resulted in significant main effects for effectiveness, $F_{(2, 118)} = 18.99$, $p < 0.001$, partial $\eta^2 = 0.003$. The medium intensity of gestures was the most effective style, followed by strong and smooth gestures. The *post hoc* tests showed significant differences among all the conditions. There also were significant differences for attractiveness, $F_{(2, 118)} = 4.63$, $p = 0.010$, partial $\eta^2 = 0.001$. The most attractive style was medium gestures, followed by strong and smooth. There were no significant differences between medium and strong intensities of gestures, according to the *post hoc* tests.

The interaction between voice pitch and gestures also was significant in effectiveness, $F_{(4, 116)} = 17.62$, $p < 0.001$, partial $\eta^2 = 0.005$. The medium variations with medium intensity of gestures were the most effective style, followed by medium pitch variations with strong and smooth gestures. The second group was formed by the strategies with medium pitch variations and the combination of medium, smooth, and strong gestures, in that order. Finally, the fewer effective styles were composed of the strategies with few pitch variations combined with strong, medium, and smooth intensities of gestures, in that order.

For attractiveness, the interaction also was significant, $F_{(4, 116)} = 27.43$, $p < 0.001$, partial $\eta^2 = 0.008$. The most attractive strategies, as in effectiveness, were the styles with medium pitch variations combined, in this order, with strong, medium, and smooth intensities of gestures. The next styles were with many pitch variations and, in this order, medium, strong, and smooth gestures. Finally, the less attractive were the strategies with few pitch variations and, in this order, strong, smooth, and medium intensities of gestures. **Table 1** shows the descriptive statistics for these two variables. These results confirm H1 and H3.

TABLE 1 | Effectiveness and attractiveness of voice pitch and gestures.

Variable	Voice pitch	Hand gestures	*Mean*	*SD*
Effectiveness	Few	Smooth	1.76	0.77
		Medium	1.78	0.75
		Strong	1.82	0.84
		Total	1.78	0.78
	Medium	Smooth	2.69	1.10
		Medium	2.99	1.15
		Strong	2.98	1.13
		Total	2.89	1.14
	Many	Smooth	2.58	1.01
		Medium	2.61	1.07
		Strong	2.44	0.94
		Total	2.55	1.02
Attractiveness	Few	Smooth	1.63	0.83
		Medium	1.58	0.76
		Strong	1.66	0.79
		Total	1.62	0.79
	Medium	Smooth	2.67	1.19
		Medium	2.94	1.21
		Strong	2.98	1.18
		Total	2.87	1.20
	Many	Smooth	2.37	1.18
		Medium	2.47	1.22
		Strong	2.43	1.16
		Total	2.43	1.19

Arousal and Valence

To test H2 and H4, a 3-pitch variation (few–medium–many) by 3 intensities of gestures (smooth–medium–strong) by 2 scripts (discourse one and discourse two) factorial MANOVA was performed on self-reported arousal and valence. As in the other variables, there were no significant differences by script.

In self-report, voice pitch had a significant main effect on arousal, $F_{(2, 118)} = 1{,}507$, $p < 0.001$, partial $\eta^2 = 0.178$. The medium strategy elicited the highest arousal, followed by the many and the few pitch variations. The *post hoc* tests indicated significant differences among all the conditions. Concerning valence, there were significant differences for voice pitch, $F_{(2, 118)} = 2{,}273$, $p < 0.001$, partial $\eta^2 = 0.246$. The strategy with many pitch variations was perceived as the most positive, followed by the medium. The strategy with few changes was considered negative. The *post hoc* tests showed significant differences among all the conditions.

The intensity of gestures had a significant main effect on arousal, $F_{(2, 118)} = 4.20$, $p = 0.015$, partial $\eta^2 = 0.001$. The medium intensity of gestures achieved the highest arousal, followed by strong and smooth gestures. The *post hoc* tests showed no significant differences between the medium and strong gestures. In valence, there also were significant differences, $F_{(2, 118)} = 3.23$, $p = 0.019$, partial $\eta^2 = 0.000$. The medium intensity of gestures was perceived as less negative, followed by strong and smooth gestures. There were no significant differences between the medium and strong styles, as shown in the *post hoc* tests.

The interaction between voice pitch and gestures also was significant in arousal, $F_{(2, 118)} = 24.74$, $p < 0.001$, partial $\eta^2 = 0.007$, and valence, $F_{(2, 118)} = 18.81$, $p < 0.001$, partial $\eta^2 = 0.005$. The stronger arousal was stimulated by combining the strategies

TABLE 2 | Self-reported arousal and valence of voice pitch and hand gestures.

Variable	Voice pitch	Hand gestures	*Mean*	*SD*
Arousal	Few	Smooth	1.54	0.84
		Medium	1.45	0.76
		Strong	1.57	0.85
		Total	1.52	0.82
	Medium	Smooth	2.54	1.17
		Medium	2.78	1.23
		Strong	2.85	1.26
		Total	2.73	1.23
	Many	Smooth	2.56	1.15
		Medium	2.58	1.18
		Strong	2.38	1.09
		Total	2.51	1.15
Valence	Few	Smooth	−2.07	1.92
		Medium	−2.16	1.88
		Strong	−2.04	1.95
		Total	−2.10	1.92
	Medium	Smooth	0.31	2.37
		Medium	0.77	2.55
		Strong	0.87	2.47
		Total	0.65	2.48
	Many	Smooth	0.97	2.17
		Medium	0.96	2.27
		Strong	0.60	2.27
		Total	0.86	2.25

with medium pitch variations and strong gestures, followed by medium and smooth. The second group was composed of the styles with many pitch variations combined with medium, smooth, and strong gestures, in that order. The third group was a few pitch variations combined, in this order, with strong, smooth, and medium intensities of gestures. The most positive valence was elicited by many pitch variations combined with smooth intensity of gestures followed by many pitch variations-medium intensity of gestures, medium-strong, medium-medium, many-strong, medium-smooth. The most negative strategies were a few pitch variations combined with, in this order, smooth, strong, and medium intensities. **Table 2** shows the results.

Physiological arousal was analyzed using a 3-pitch variation (few–medium–many) ANOVA. The results were significant, $F_{(2, 118)} = 15.43$, $p < 0.001$, partial $\eta^2 = 0.012$. The medium style achieved the highest arousal ($M = 0.43$; $SD = 0.97$), followed by many ($M = 0.28$; $SD = 0.60$), and few ($M = 0.23$; $SD = 0.54$). The strategy with medium pitch variations stimulated the strongest emotional activation followed by many and few changes. There were no significant results among many and few.

About gestures, the data were analyzed using a 3-intensity (smooth–medium–strong) ANOVA. There were no significant results, $F_{(2, 118)} = 2.29$, $p = 0.101$, partial $\eta^2 = 0.002$. The strategy with strong gestures achieved the greater arousal ($M = 0.43$; $SD = 0.81$), followed by medium ($M = 0.34$; $SD = 0.91$), and smooth gestures ($M = 0.28$; $SD = 0.64$).

However, the interaction between voice pitch and gestures in arousal was significant, $F_{(4, 116)} = 2.54$, $p = 0.038$, partial $\eta^2 = 0.004$. The combination with the strongest activation was medium pitch variations-strong gestures, followed by medium variations-medium gestures, medium variations-smooth gestures, many variations-medium gestures, many variations-smooth gestures, many variations-strong gestures, few variation-medium gestures, few variations-strong gestures, and few variations-smooth gestures. **Figure 3** shows the results.

Finally, in valence, there were significant results for voice pitch measured with the facial recognition of emotions (Facereader), $F_{(2, 118)} = 256.75$, $p < 0.001$, partial $\eta^2 = 0.324$. The most positive style was many pitch variations, followed by medium and few changes. There were no significant differences between medium and many variations, as shown in the *post hoc* tests. Along with this, there were significant differences in happiness, $F_{(2, 118)} = 220.52$, $p < 0.001$, partial $\eta^2 = 0.292$, and sadness, $F_{(2, 118)} = 113.87$, $p < 0.001$, partial $\eta^2 = 0.175$. The happiest strategy was with many pitch variations, followed by medium and few changes, whereas the saddest style was the opposite, in this order: few, medium, and many. *Post hoc* tests showed significant differences among the three styles in happiness and sadness.

The intensity of gestures also had a significant main effect on valence, $F_{(2, 118)} = 97.09$, $p < 0.001$, partial $\eta^2 = 0.153$. The most positive style was the medium intensity, followed by strong and smooth. There were significant differences among the three styles, as shown in the *post hoc* tests. Also, there were significant differences in the two main emotions, happiness $F_{(2, 118)} = 9.03$, $p < 0.001$, partial $\eta^2 = 0.017$, and sadness, $F_{(2, 118)} = 77.94$, $p < 0.001$, partial $\eta^2 = 0.127$. Participants showed more happiness with the medium intensity of gestures, followed by strong and smooth. They expressed more sadness with the style with smooth intensity of gestures followed by medium and strong. There were no significant differences between medium and strong gestures, as shown in the *post hoc* tests.

The interaction between voice pitch and gestures also was significant in valence, $F_{(4, 116)} = 55.91$, $p < 0.001$, partial $\eta^2 = 0.173$, happiness, $F_{(4, 116)} = 16.94$, $p < 0.001$, partial $\eta^2 = 0.060$, and sadness, $F_{(4, 116)} = 19.53$, $p < 0.001$, partial $\eta^2 = 0.068$. The most positive strategy was the medium followed by all the styles with many pitch variations combined with medium, strong, and smooth intensities of gestures, in that order. However, participants expressed more happiness with the strategies of many pitch variations with medium, strong, and smooth intensities of gestures, in this order. The less happy strategy combined a few pitch variations with smooth gestures. In contrast, the strategy that elicited more sadness was a few pitch variations with smooth gestures followed by medium-strong. The less sad strategy was medium pitch variations with strong gestures. These results confirm H2 and H4. **Table 3** shows the descriptive statistics.

DISCUSSION

This study aimed to analyze the effectiveness, attractiveness, and emotional response (arousal and valence) of different pitch

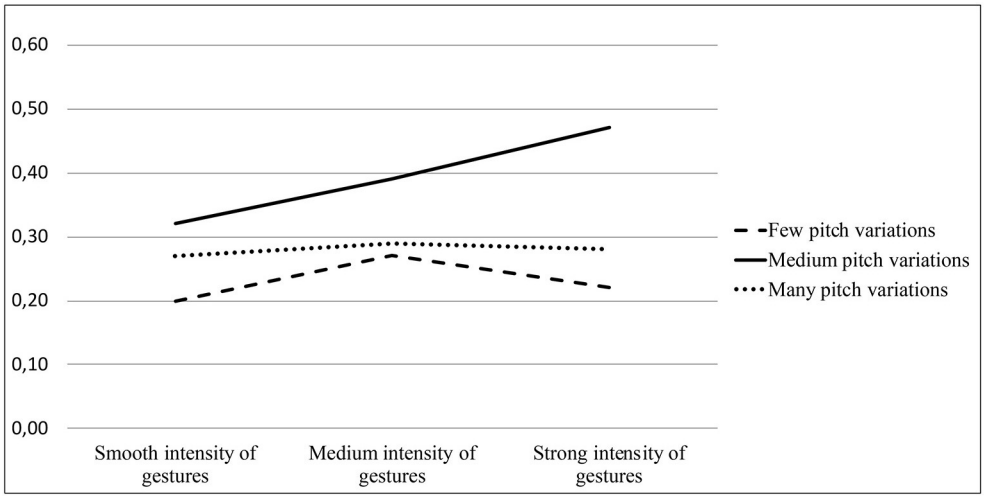


FIGURE 3 | EDA interaction voice pitch-hand gestures.

variations and intensities of hand gestures in public discourses. To do so, this research combined different methods: self-report, physiological response, and facial recognition of emotions, with a within-subject design, as it is common in physiological studies due to their complexity. The findings showed that the non-verbal communication cues examined in this research were relevant in determining how effective and attractive a public speech is and affecting the participants' emotional response.

About effectiveness and attractiveness, the results indicated that the strategy with medium pitch variations was considered the most effective and attractive, followed by the styles with many and few changes. This result was expected, in line with previous studies that revealed better perception of pitch variations (Rodero et al., 2017), significantly when they are moderated (Rodero et al., 2022) and compared to monotonous strategies (Burgoon et al., 1990; Glass, 1991). A balanced strategy with pitch changes, nor few nor excessive, can be perceived as more dynamic (Addington, 1968) and, consequently, more effective and attractive, as the data have shown. The second-best style was many pitch variations, as a more expressive voice is always better perceived than dull strategies (Rodero et al., 2022). Therefore, the less effective and attractive style was few variations, as it sounded more monotonous, and this characteristic hindered perception (Knapp and Hall, 2007; Strangert and Gustafson, 2008).

As with voice pitch, in gestures, the data showed the same trend for perception. The most effective and attractive style was the medium intensity of gestures, followed by strong and smooth. A balanced strategy also was the best perceived by the participants. The medium intensity of gestures was enough to be rated as the most effective and attractive, in line with previous research in which speakers making hand gestures have been assessed with positive traits, such as effective, persuasive, credible, dominant, extrovert, sociable, and honest (Burgoon et al., 1990; Beattie and Shovelton, 2005; Maricchiolo et al., 2009; Neff et al., 2010; Jackob et al., 2011; Gnisci and Pace, 2014; Peters and Hoetjes, 2017; King et al., 2020).

The interaction between these two non-verbal features revealed a slightly different pattern. The most effective and attractive strategies were those with medium pitch variations combined with medium (effective) and strong gestures (attractive). Therefore, the best styles were medium pitch variations with medium or strong gestures. When both features act together, the expressivity level is the same as individually for effectiveness (medium-medium) but higher (medium-strong) for attractiveness. These results are according to the studies showing that there is a positive perception when speakers use pitch variations (Hincks, 2004; Strangert, 2005; Yang et al., 2020; Rodero et al., 2022) and a high quantity of gestures (Maricchiolo et al., 2009; Gnisci and Pace, 2014; Van Edwards, 2017), mainly to be evaluated as attractive, as the data show.

Concerning self-reported arousal, both in voice pitch and in gestures, the medium strategies (medium pitch variations and medium intensity of gestures) achieved the highest level. These data also represent an expected result, in line with previous studies. If a speaker talks with enough pitch changes and hand movements, these variations can increase arousal. A balanced strategy with medium pitch variations, guided by the Principle of Distinctive and Contrastive Coherence of Prosody (Rodero, 2015), elicited the strongest self-report arousal. The second-best strategy was the styles with many changes that, although achieved a high activation, were less intense than the medium ones. We can suggest that these variations and intensity of gestures were excessive, and therefore, counter-productive. Finally, the worst evaluated strategies were the less expressive styles: few pitch variations with smooth intensity of gestures, as it seems reasonable. The lack of change with a more regular voice and smooth physical hand movements reduced

TABLE 3 | Valence, happiness, and sadness of voice pitch and hand gestures.

Variable	Voice pitch	Hand gestures	*Mean*	*SD*
Valence	Few	Smooth	−0.38	0.27
		Medium	0.14	0.21
		Strong	−0.009	0.30
		Total	−0.07	0.34
	Medium	Smooth	0.20	0.10
		Medium	0.31	1.14
		Strong	0.15	0.25
		Total	0.22	0.19
	Many	Smooth	0.23	0.23
		Medium	0.27	0.17
		Strong	0.22	0.14
		Total	0.24	0.19
Happiness	Few	Smooth	0.13	0.10
		Medium	0.12	0.08
		Strong	0.04	0.05
		Total	0.10	0.09
	Medium	Smooth	0.15	0.11
		Medium	0.39	0.32
		Strong	0.35	0.34
		Total	0.29	0.29
	Many	Smooth	0.47	0.30
		Medium	0.46	0.26
		Strong	0.45	0.23
		Total	0.46	0.27
Sadness	Few	Smooth	0.90	0.17
		Medium	0.54	0.35
		Strong	0.68	0.28
		Total	0.71	0.31
	Medium	Smooth	0.67	0.35
		Medium	0.40	0.27
		Strong	0.34	0.27
		Total	0.47	0.33
	Many	Smooth	0.40	0.17
		Medium	0.39	0.18
		Strong	0.35	0.18
		Total	0.38	0.18

arousal, according to the "sensory adaptation" (Wark et al., 2007).

However, in physiological arousal, the results were different in gestures. The moderate pitch variations had to be complemented with strong gestures to achieve the most significant body activation. The acoustic changes were not enough and needed intense hand movements in gestures. This result is in line with the data on the interaction between the two non-verbal features. In this case, the strongest arousal and activation in self-report and physiological response was the combination of the strategies with medium pitch variations and strong gestures, followed by medium and smooth. Therefore, for physiological arousal in gestures and when both features act together, voice pitch is complemented with strong gestures to attain the highest arousal level. This result suggests that the expressivity attained by voice pitch was not enough to aroused participants, and it had to be completed with a strong intensity of gestures to achieve an adequate expressivity level. Participants perceived this combination as the most attractive and were more aroused and activated when exposed to it.

Finally, there were no differences between voice pitch and gestures regarding self-reported valence and facial recognition of emotions (Facereader). Here, the results are different than in other variables. The strategy with many pitch variations was perceived as the most positive in voice pitch, followed by medium and few changes. The style with few changes was considered the most negative. In gestures, the most positive was medium, followed by strong and smooth gestures, both in self-report and in facial recognition. Therefore, many pitch changes were needed to be perceived as positive, while in gestures, the medium intensity was enough.

However, as in the other variables, the results slightly differed in the interaction between the two non-verbal features. In this case, the most positive strategy was the medium styles in both variables, followed by the strategies with many pitch variations combined with medium, strong, and smooth gestures. So, although in arousal, a higher intensity of gestures was necessary, the results revealed that the balanced or moderated strategy was the most positive, once again as the high change styles could have been considered excessive. On the contrary, the most negative styles were the less expressive: few pitch variations and smooth intensity of gestures. As the data showed, this monotony and lack of movement could have affected participants' attention and, thus, arousal and valence.

Finally, the happiest strategy was for many pitch variations, followed by medium and few changes, whereas the saddest style was the opposite. In gestures, the results were the same (strong–medium–smooth) in both variables. As it seems logical, participants expressed the highest happiness level (shown in their facial expressions) with the strategies with many pitch variations and strong intensity of gestures. In contrast, sadness was shown especially with the fewer expressivity strategies (few pitch variations and smooth gestures).

In conclusion, the results indicated that participants perceived as more effective, attractive, and positive and felt happier and more aroused with medium-level strategies (medium pitch variations and medium intensity of gestures) when both features were analyzed separately. This balanced strategy contained enough pitch variations and hand movements to be positively perceived and elicit a beneficial emotional response in arousal and valence. However, when the analyzed features acted together, the medium pitch variation strategy had to be complemented with strong intensity of gestures to be perceived as the most attractive and achieve the highest arousal. This result can indicate that the strategy with medium pitch variations in this study could not have been sufficiently expressive to raise high attractiveness and arousal levels. In any case, the results are

consistent individually, and in the rest of the variables, so it can be concluded that the best combination between voice pitch and gestures is an expressive style medium-medium or medium-strong, but not many-strong, as it can be excessive, or few-smooth, as it is monotonous.

LIMITATIONS AND FUTURE RESEARCH

This study has three main limitations that must be considered to interpret the results. The first limitation was to record all the presentations controlling voice pitch and gestures variations while the rest of the variables maintained stable. When someone speaks, many variables are acting at the same time. Despite this difficulty, the training hours with the actors and the posterior analysis confirm that the results were accurate. Along with this, the second limitation was the limited number of analyzed strategies. Further research should extend the results to different voice and gestures combinations and other features. Finally, the third limitation was the number of actors (four) and the short speeches. We hypothesize that, in this study, the style of these concrete actors could have influenced the results. However, the video recording and the analysis showed that all of them were neutral in their performances and achieved the described styles according to the previously given guidelines. Anyhow, it would be convenient to extend the number of speakers by including more ethnicities and appearances. Also, future research should analyze long discourses. In this study, the duration of the speeches was deliberately short due to the extensive number of data generated, particularly by the physiological analysis. Finally, participants watched a high number of videos with the same content. As the order and the speakers were varied, and each session was completed in 40 min, signs of tiredness were not perceived, but a slight fatigue effect could have influenced the results.

All in all, the contributions of this study were both scientific and professional. The results of this study allow us to advance in the analysis of non-verbal cues, such as voice pitch and gestures and their influence on emotional response. Along with this, the findings can be applied to train speakers in public discourses. The main recommendation is to use an expressive performance with pitch variations and medium or strong intensity of hand gestures.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary materials, further inquiries can be directed to the corresponding author/s.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Barcelona School of Management. The patients/participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

AUTHOR CONTRIBUTIONS

ER states that her contribution as an author of the present manuscript follows the authorship definition of the International Committee of Medical Journal Editors. She has directed the research and done all the contributions to the conception or design of the work. She was in charge of the project coordination, acquisition, analysis, and interpretation of data for the work. Furthermore, she has revised the work critically for important intellectual content and has prepared the final version to be published. Finally, she agrees to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

FUNDING

This research was funded by the Barcelona School of Management.

REFERENCES

Addington, D. W. (1968). The relationship of selected vocal characteristics to personality perception. *Speech Monogr.* 35, 492–503. doi: 10.1080/03637756809375599

Bar, M., Neta, M., and Linz, H. (2006). Very first impressions. *Emotion* 6, 269–278. doi: 10.1037/1528-3542.6.2.269

Beattie, G., and Shovelton, H. (2005). Why the spontaneous images created by the hands during talk can help make TV advertisements more effective. *Br. J. Psychol.* 96, 21–37. doi: 10.1348/000712605X103500

Belin, P., Fecteau, S., and Bédard, C. (2004). Thinking the voice: neural correlates of voice perception. *Trend Cogn. Sci.* 8, 129–135. doi: 10.1016/j.tics.2004.01.008

Berger, K. W., and Popelka, G. R. (1971). Extra-facial gestures in relation to speechreading. *J. Commun. Disord.* 3, 302–308. doi: 10.1016/0021-9924(71)90036-0

Birdwhistell, R. (1970). *Kinesics and Context*. Philadelphie, PA: University of Pennsylvania Press. doi: 10.9783/9780812201284

Boersma, P., and Weenink, D. (2022). Praat: doing phonetics by computer. 10. Available online at: https://www.fon.hum.uva.nl/praat/

Bradley, M. M., and Lang, P. J. (1994). Measuring emotion: the self-assessment manikin and the semantic differential. *J. Behav. Ther. Exper. Psychiat.* 25, 49–59. doi: 10.1016/0005-7916(94)90063-9

Burgoon, J. K., Birk, T., and Pfau, M. (1990). Nonverbal behaviors, persuasion, and credibility. *Hum. Commun. Res.* 17, 140–169. doi: 10.1111/j.1468-2958.1990.tb00229.x

Burgoon, J. K., Guerrero, L. K., and Floyd, K. (2010). *Nonverbal Communication*. Routledge.

Chebat, J. C., El Hedhli, K., Gélinas-Chebat, C., and Boivin, R. (2007). Voice and persuasion in a banking telemarketing context. *Percep. Motor Skill.* 104, 419–437. doi: 10.2466/pms.104.2.419-437

Cherry, E. C. (1953). Some experiments on the recognition of speech, with one and two ears. *J. Acoust. Soc. Am.* 25, 975–979. doi: 10.1121/1.1907229

Clark, T., and Greatbatch, D. (2011). Audience perceptions of charismatic and non-charismatic oratory: The case of management gurus. *Leadersh. Q.* 22:22–32. doi: 10.1016/j.leaqua.2010.12.004

Clough, S., and Duff, M. C. (2020). The role of gesture in communication and cognition: implications for understanding and treating

neurogenic communication disorders. *Front. Hum. Neurosci.* 14, 323. doi: 10.3389/fnhum.2020.00323

Dargue, N., Sweller, N., and Jones, M. P. (2019). When our hands help us understand: a meta-analysis into the effects of gesture on comprehension. *Psychol. Bull.* 145, 765–784. doi: 10.1037/bul0000202

Dawson, M. E., Schell, A. M., and Filion, D. L. (2000). "The electrodermal system", in Cacioppo J., Tassinary L. G., Berntson G. G., eds. *Handbook of Psychophysiology*. New York, NY: Cambridge University Press. p. 200–223.

Ekman, P., and Friesen, W. V. (1969). The repertoire of nonverbal behavior. *Semiotica* 1, 49–98. doi: 10.1515/semi.1969.1.1.49

Ekman, P., Friesen, W. V., O'Sullivan, M., and Scherer, K. (1980). Relative importance of face, body, and speech in judgments of personality and affect. *J. Person. Soc. Psychol.* 38, 270–277. doi: 10.1037/0022-3514.38.2.270

Elbert, S. P., and Dijkstra, A. (2014). An experimental test of the relationship between voice intonation and persuasion in the domain of health. *Psychol. Health.* 29, 1014–1031. doi: 10.1080/08870446.2014.903482

Gélinas-Chebat, C., Chebat, J. C., and Vaninsky, A. (1996). Voice and advertising: Effects of intonation and intensity of voice on source credibility attitudes toward the advertised service and the intent to buy. *Percept. Motor Skill.* 85, 243–262. doi: 10.2466/pms.1996.83.1.243

Glass, L. (1991). *Say it Right*. How to talk in any social situation. HarperAudio.

Gnisci, A., and Pace, A. (2014). "The Effects of Hand Gestures on Psychosocial Perception: A Preliminary Study", in *Recent Advances of Neural Network Models and Applications*. S. Bassis, A. Esposito, and F. C. Morabito (Eds.), Smart Innovation, Systems and Technologies (Vol. 26, pp. 305–314). Springer. doi: 10.1007/978-3-319-04129-2_30

Gordon, R. A., Druckman, D., Rozelle, R. M., and Baxter, J. C. (2006). *Non-Verbal Behaviour as Communication: Approaches, Issues and Research.* The Handbook of Communication Skills. Routledge.

Graham, J. A., and Argyle, M. (1975). A cross-cultural study of the communication of extra-verbal meaning by gestures (1). *Int. J. Psychol.* 10, 57–67. doi: 10.1080/00207597508247319

Guyer, J. J., Briñol, P., Petty, R. E., and Horcajo, J. (2019). Nonverbal behavior of persuasive sources: a multiple process analysis. *J. Nonverbal Behav*. 43, 203–231. doi: 10.1007/s10919-018-00291-x

Harris, T. (2003). Listening with your eyes: the importance of speech-related gestures in the language classroom. *Fore. Langu. Ann.* 36, 180–187. doi: 10.1111/j.1944-9720.2003.tb01468.x

Hill, K. T., and Miller, L. M. (2010). Auditory attentional control and selection during cocktail party listening. *Cerebr. Cortex* 20, 583–590. doi: 10.1093/cercor/bhp124

Hincks, R. (2004). "Processing the prosody of oral presentations", in *ICALL-2004, Symposium on Computer Assisted Learning.*

Hirschberg, J., and Pierrehumbert, J. (1986). "Intonational structuring of discourse", in *Proceedings of 24th Meeting of the Association for Computational Linguistics. Nueva York* p. 136–144. doi: 10.3115/981131.981152

Holle, H., and Gunter, T. C. (2007). The role of iconic gestures in speech disambiguation: ERP evidence. *J. Cogn. Neurosci.* 19, 1175–1192. doi: 10.1162/jocn.2007.19.7.1175

Hostetter, A. B. (2011). When do gestures communicate? A meta-analysis. *Psychol. Bull.* 137, 297–315. doi: 10.1037/a0022128

Ianì, F., and Bucciarelli, M. (2017). Mechanisms underlying the beneficial effect of a speaker's gestures on the listener. *J. Memory Lang.* 96, 110–121. doi: 10.1016/j.jml.2017.05.004

Iverson, J. M., and Goldin-Meadow, S. (1998). Why people gesture when they speak. *Nature* 396, 228. doi: 10.1038/24300

Jackob, N., Roessing, T., and Petersen, T. (2011). The effects of verbal and nonverbal elements in persuasive communication: Findings from two multi-method experiments. *Communications.* 36, 245–271. doi: 10.1515/comm.2011.012

Jannedy, S., and Mendoza-Denton, N. (2005). "Structuring information through gesture and intonation", in: *Interdisciplinary studies on information structure: ISIS; working papers of the SFB* 632, 199–244.

Kelly, S., Healey, M., Özyürek, A., and Holler, J. (2015). The processing of speech, gesture, and action during language comprehension. *Psychon. Bull. Rev.* 22, 517–523. doi: 10.3758/s13423-014-0681-7

Kelly, S. D., Özyürek, A., and Maris, E. (2010). Two sides of the same coin: Speech and gesture mutually interact to enhance comprehension. *Psychol. Sci.* 21, 260–267. doi: 10.1177/0956797609357327

Kendon, A. (1981). Geography of gesture. *Semiotica.* 37, 129–163.

King, J. P. J., Loy, J. E., Rohde, H., and Corley, M. (2020). Interpreting nonverbal cues to deception in real-time. *PLoS ONE* 15, 1–25. doi: 10.1371/journal.pone.0229486

Knapp, M. L. (1980). *Essentials of Nonverbal Communication*. Harcourt School.

Knapp, M. L., and Hall, J. A. (2007). *Nonverbal Communication in Human Interaction*. Thomas Learning.

Krauss, R. M., Morrel-Samuels, P., and Colasante, C. (1991). Do conversational hand gestures communicate? *J. Person. Soc. Psychol.* 61, 743. doi: 10.1037/0022-3514.61.5.743

Lang, P. J., and Bradley, M. M. (2013). Appetitive and defensive motivation: goal-directed or goal-determined? *Emot. Rev.* 5, 230–234. doi: 10.1177/1754073913477511

Levi, S. V., and Pisoni, D. B. (2007). "Indexical and linguistic channels in speech perception: Some effects of voiceovers on advertising outcomes", in T.M. Lowrey (Ed.), *Psycholinguistic Phenomena in Marketing Communications.* Lawrence Erlbaum Associates.

Loehr, D. P. (2004). Gesture and intonation. Washington: Georgetown University.

Maricchiolo, F., Gnisci, A., Bonaiuto, M., and Ficca, G. (2009). Effects of different types of hand gestures in persuasive speech on receivers' evaluations. *Langu. Cogn. Process.* 24, 239–266. doi: 10.1080/01690960802159929

Neff, M., Wang, Y., Abbott, R., and Walker, M. (2010). "Evaluating the effect of gesture and language on personality perception in conversational agents", in: *International Conference on Intelligent Virtual Agents*. (pp. 222-235). Springer, Berlin, Heidelberg. doi: 10.1007/978-3-642-15892-6_24

Newman, R., Furnham, A., Weis, L., Gee, M., Cardos, R., Lay, A., et al. (2016). Nonverbal presence: how changing your behaviour can increase your ratings for persuasion, leadership and confidence. *Psychology* 07, 488–499. doi: 10.4236/psych.2016.74050

Niebuhr, O., and Silber-Varod, V. (2021). How versatility performance influences perception of charismatic speech: a study on two Israeli politicians. *Interact. Stud.* 22, 303–342. doi: 10.1075/is.00007.nie

Peters, J., and Hoetjes, M. W. (2017). "The effect of gesture on persuasive speech", in *Interspeech* 2017, p. 659–663. doi: 10.21437/Interspeech.2017-194

Potter, R. F., and Bolls, P. D. (2012). *Psychophysiological Measurement and Meaning: Cognitive and Emotional Processing of Media.* Routledge. doi: 10.4324/9780203181027

Riseborough, M. G. (1981). Physiographic gestures as decoding facilitators: Three experiments exploring a neglected facet of communication. *J. Nonverbal Behav.* 5, 172–183. doi: 10.1007/BF00986134

Rodero, E. (2013). The perception of a broadcasting voice. *US-China Educ. Rev. A and B.* 3, 225–230.

Rodero, E. (2015). The principle of distinctive and contrastive coherence of prosody in radio news: an analysis of perception and recognition. *J. Nonverbal Behav.* 39, 79–92. doi: 10.1007/s10919-014-0201-5

Rodero, E. (2021). "*Measuring People's Cognitive and Emotional Response*", Psychophysiological methods to study sound design, in Filimowicz, M. (ed.), Doing Research in Sound Design. New York, NY: Routledge. doi: 10.4324/9780429356360-5

Rodero, E., Larrea, O., Rodríguez-de-Dios, I., and Lucas, I. (2022). The expressive balance effect. perception and physiological responses of prosody and gestures. *J. Langu. Soc. Psychol.* doi: 10.1177/0261927X221078317. [Epub ahead of print].

Rodero, E., Larrea, O., and Vázquez, M. (2013). Male and female voices in commercials. Analysis of effectiveness, adequacy for product, attention and recall. *Sex Roles* 68, 349–362. doi: 10.1007/s11199-012-0247-y

Rodero, E., Potter, R. F., and Prieto, P. (2017). Pitch range variations improve cognitive processing of audio messages. *Hum. Commun. Res.* 43, 397–413. doi: 10.1111/hcre.12109

Rodero, E., Mas, L., and Blanco, M. (2019). The influence of prosody on politician's credibility. *J. Appl. Linguist. Profess. Pract.* 11, 89–111. doi: 10.1558/japl.32411

Rosenberg, A., and Hirschberg, J. (2009). Charisma perception from text and speech. *Speech Commun.* 51, 640–655. doi: 10.1016/j.specom.2008.11.001

Scott, W. (1920). *Lessons in Elocution, or, a Selection of Pieces, in Prose and Verse, for the Improvement of Youth in Reading and Speaking*. Montpelier: E. P. Walton.

Strangert, E. (2005). "Prosody in public speech: analyses of a news announcement and a Political interview", in *INTERSPEECH-2005, Eurospeech, 9th European Conference on Speech Communication and Technology*. doi: 10.21437/Interspeech.2005-827

Strangert, E., and Gustafson, J. (2008). "What makes a good speaker? subject ratings, acoustic measurements and perceptual evaluations", in: *Ninth Annual Conference of the International Speech Communication Association*. p. 1688–1691. doi: 10.21437/Interspeech.2008-368

Talley, L., and Temple, S. (2015). How leaders influence followers through the use of nonverbal communication. *Leaders. Organiz. Develop. J.* 36, 69–80. doi: 10.1108/LODJ-07-2013-0107

Tench, P. (2015). *The Intonation Systems of English*. Bloomsbury Publishing.

Van Edwards, V. (2017). You are contagious. Available online at: https://www.youtube.com/watch?v=cef35Fk7YD8

Varghese, A. L., and Nilsen, E. S. (2020). Is that how you should talk to her? Using appropriate prosody affects adults', but not children's, judgments of communicators' competence. *J. Lang. Soc. Psychol.* 39, 738–750. doi: 10.1177/0261927X19871692

Wagner, P., Malisz, Z., and Kopp, S. (2014). Gesture and speech in interaction: an overview. *Speech Commun.* 57, 209–232. doi: 10.1016/j.specom.2013.09.008

Wark, B., Lundstrom, B. N., and Fairhall, A. (2007). Sensory adaptation. *Curr. Opin. Neurobiol.* 17 , 423–429. doi: 10.1016/j.conb.2007.07.001

Weinstein, N., Vansteenkiste, M., and Paulmann, S. (2020). Don't you say it that way! Experimental evidence that controlling voices elicit defiance. *J. Exper. Soc. Psychol.* 88, 103949. doi: 10.1016/j.jesp.2019.103949

Wells, J. C. (2007). *English Intonation*. Cambridge: Cambridge University Press.

Yang, Z., Huynh, J., Tabata, R., Cestero, N., Aharoni, T., and Hirschberg, J. (2020). "What makes a speaker charismatic? Producing and perceiving charismatic speech", in *Procrrdings 10th International Conference on Speech Prosody 2020*. pp. 685–689. doi: 10.21437/SpeechProsody.2020-140

Conflict of Interest: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Rodero. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers | Frontiers in Psychology

BRIEF RESEARCH REPORT
published: 09 June 2022
doi: 10.3389/fpsyg.2022.861366

Research on the Visual Imagery of Posters Based on the Culture Code Theory of Design

Luo Wen†, Wang Jingjing†, Wang Chen and Sun Luyu*

School of Design, South China University of Technology, Guangzhou, China

According to the Culture Code Theory of Design, posters can be divided into strategy aspect, meaning aspect, and technical aspect. This study explored the impact of visual imagery on poster communication and effect focus on the meaning and technical aspects which directly related to the audiences. The visual imagery preference of ten public poster samples is collected from 40 participants by online questionnaire with the semantic differential method. The results show that the audience's visual imagery at the technical aspect is mainly related to the richness of colors, graphics, and texts. When receiving the information from the meaning aspect of the poster correctly, the visual imagery will be influenced by the type of theme. Most importantly, the same or similar visual imagery at the technical and meaning aspect encourages audiences to explore the deep information of posters. Poster works that keep visual imagery coherent and smooth at the two aspects can achieve a better effect.

Keywords: visual imagery, culture code theory of design, poster design, visual communication, semantic differential method

OPEN ACCESS

Edited by:
Francesca D'Errico,
University of Bari Aldo Moro, Italy

Reviewed by:
Larissa D'Angelo,
University of Bergamo, Italy
Jie Zhang,
Jiangnan University, China

****Correspondence:***
Wang Chen
wangchen@scut.edu.cn

†These authors share first authorship

Specialty section:
This article was submitted to Emotion Science, a section of the journal Frontiers in Psychology

Received: *24 January 2022*
Accepted: *25 March 2022*
Published: *09 June 2022*

Citation:
Wen L, Jingjing W, Chen W and Luyu S (2022) Research on the Visual Imagery of Posters Based on the Culture Code Theory of Design. Front. Psychol. 13:861366. doi: 10.3389/fpsyg.2022.861366

INTRODUCTION

Imagery, also known as psychological image, is the psychological feeling based on physical image and personal perception. Imagery can be both produced in the original consciousness field of the information producer, and extracted and formed after the audience receives the information (Tao, 2018). According to the source of sensation, imagery includes visual, auditory, olfactory, taste, and tactile imagery, among which visual imagery plays the most important role (Wu and Liu, 2016). The art psychology expert Rudolf Arnheim has proved from a large number of facts that whenever we think about a problem, there will be some imagery as the starting point or basis. In his view, creative thinking is inseparable from visual imagery in any cognitive field (Science and Art), and visual imagery is the bridge connecting science and art, rationality and sensibility (Zhencheng, 2005).

The poster is a functional visual design, which plays an important role in information communication and owns a significant relationship with visual imagery (Purnomo and Hidayatuloh, 2020). The visual imagery connection between designers and audiences influences the information exchange. According to the purpose, posters can be divided into four categories: public, political, cultural, and commercial posters. The main purpose of public posters is to spread the values of truth, goodness, and beauty to the public and serve the establishment of social morality. The main purpose of political posters is to spread political ideology. Cultural posters serve all

kinds of cultural products and cultural exhibitions. Commercial posters are for-profit and serve enterprises and businesses (Yang, 2016). Compared with the other three categories, public posters have the widest audience and this research takes public posters as the main research object.

Current pieces of research mainly focus on the visual attraction and semantic understanding of posters. Most scholars believe that a good poster design needs to attract the audience's attention at first. For example, colors and graphics with strong contrast will leave a deep visual impression on the audience, stimulate them to decode the information expressed by designers, and re-code it into personal cognition (Brown et al., 2013; Sentell, 2016). They apply visual dot-probe, eye tracker, semantic difference table to explore the impact of visual elements such as color, text, and composition of posters on the audience's attention while ignoring the meaning communication of the poster (Wu and Liu, 2016; Yu and Ko, 2017; Chen et al., 2020; Molek-Kozakowska, 2020). On the other hand, although there are some research on the audience's expression and emotion changes during viewing, the main methods include observation and interview. However, the observer's subjective view easily to cause the omission and the misunderstanding of the audience's feeling (Sentell, 2016; Charest et al., 2019). A combination of qualitative and quantitative research needs to be applied to the effect of posters.

As a culture system theory, the Culture Code Theory of Design was first proposed by Yufu Yang based on semiotic theory (Yufu, 2002). The Culture Code Theory of Design decomposes posters into three aspects: strategy, meaning, and technical aspect which is an effective tool for analyzing graphic design works. As shown in **Figure 1**, the strategy aspect is similar to the story structure in an article which is the designer's ingenious thinking and is related to the designer's cultural background. The meaning aspect can be compared to the words and sentences in an article. The technical aspect, also known as the element aspect, refers to all kinds of materials such as font, graphics, and colors as the single word forming a sentence. Take a work of the 4th German ANFACHEN Poster Design Award as an example shown in **Figure 1**. The broken plastic bag, ground, water, cartoon fish (technical aspect) are used to form a painful image of fish sending out the "HELP" signal on the dry ground (strategic aspect) to expresses the concern that white pollution has made marine life miserable, and calls on the public to protect marine resources (meaning aspect). In most cases, the strategy aspect will be deliberately hidden by the designer, which is difficult to grasp completely and accurately for audiences. The technical aspect and the meaning aspect are easier for audiences to receive and understand which is also the effect that the author hopes to achieve (Yufu, 2002; Yang et al., 2012). Therefore, this research mainly focuses on the meaning aspect and technical aspect of posters.

Poster design is an art design that requires both visual attraction and readability. Applying the Culture Code Theory of Design, we can completely study the visual imagery of posters from the technical aspect (color, text, graphics, etc.) and the meaning aspect (theme) to provide designers with design and creation strategies to improve the effect of poster design.

MATERIALS AND METHODS

Poster Samples

The poster sample for this research comes from the German ANFACHEN Poster Design Award[1]. Every year, the 25 best posters will be awarded at a public exhibition with thousands of international posters. It has been held for four sessions since 2017 and involves four public themes: Woman, Tolerance, House and Living, and Water. A total of 100 winning works from the four sessions were selected as the initial sample set. An expert group composed of two design teachers and two design students coded 100 posters from the meaning aspect (theme) and the technical aspect (style, composition, graphics and text, and color) of the poster as shown in **Table 1** (Zhang, 2015; Yang, 2016). Four feature sample sets are obtained by hierarchical clustering of SPSS (IBM SPSS Statistics 25, Armonk, NY, United States of America). The size of the posters in the four sample sets is 35, 24, 20, and 21, respectively. The final scale is 4:2:2:2 of the four clusters.

Adjective Words

The poster evaluation text is collected from relevant literature, magazines, and websites. The Weiciyun website[2] is used for word segmentation and word frequency statistics. A total of 54 adjectives are obtained. The expert group selects ten adjectives as the final evaluation basis for the poster work. Then, ten pairs of representative opposite words are chosen, including Soft-Intense, Popular-Individualized, Plain-Fashionable, Traditional-Modern, Elegant-Heroic, Serious-Lively, Cold-Enthusiastic, Static-Dynamic, Empty-Plump, Harmonious-Conflicting.

Experiment Procedure

Forty participants are recruited in total. Half of the participants come from design majors, and the other half come from non-design majors, with 50% male and 50% female to eliminate the potential influence of gender and specialty on experiment results. All participants are informed of the experiment procedure and sign informed consent. This experiment has been approved by the Ethics Review Committee of the South China University of Technology.

Using the 7-level Likert scale and regarding the two words of each adjective pair as two extreme cases, the participants use the scale to choose the degree of personal feeling between each adjective pair. The poster samples appear completely on the computer screen or mobile screen at a specific time. All participants need to fill in the questionnaire twice. In the first questionnaire, the participants are reminded to complete the questionnaire for each poster sample based on first impressions. The first appearance time of the poster is 7 s according to the research on the first impression of online web pages (Kim and Fesenmaier, 2008). In the second questionnaire, the participants need to fill in the questionnaire after deep thinking and record their understanding of each poster sample so the appearance time depends on themselves.

[1] https://www.anfachenaward.de/en

[2] www.weiciyun.com

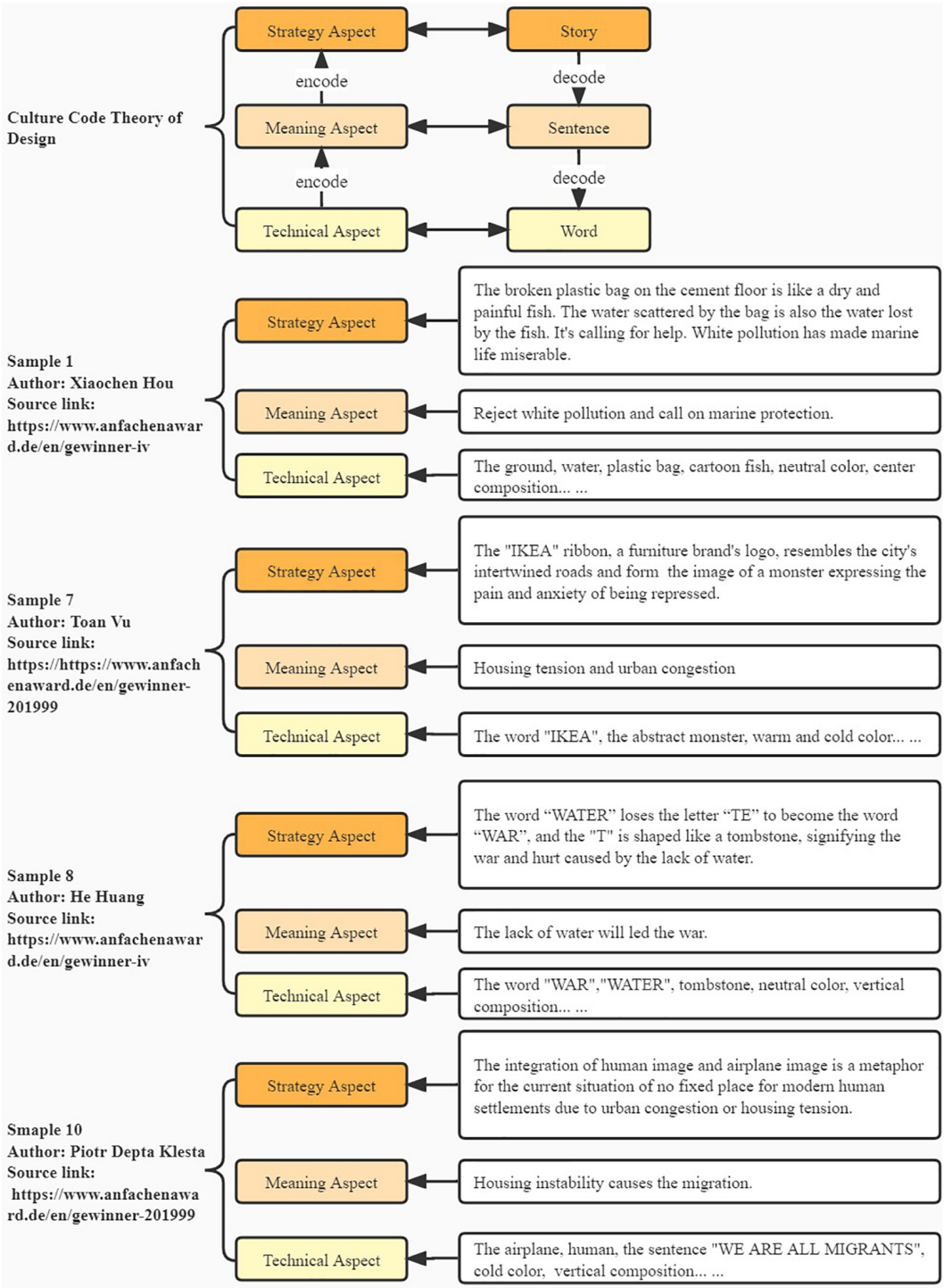


FIGURE 1 | Culture code theory of design. (Poster source: https://www.anfachenaward.de/en).

RESULTS

First, the questionnaire with the same score of ten dimensions is considered invalid and is eliminated. There is an overall analysis of the visual imagery of the poster sample at the technical aspect and the meaning aspect with the factor analysis of SPSS (IBM SPSS Statistics 25, America). The factors with large loads in multiple factors are deleted (Static-Dynamic, Empty-Plump, Harmonious-Conflicting) and three principal factors are extracted from the seven factors. The final factor analysis result is

TABLE 1 | The code rule of 10 poster samples.

Elements	Categories		
Theme	1 Woman; 4 Water	2 Tolerance;	3 House and Living;
Style	1 Abstract; 4 Cartoon;	2 Realism; 5 Text	3 Impressionism;
Composition	1 Central; 4 Vertical; 7 Surround;	2 Symmetrical; 5 Scattered; 8 Others	3 Horizontal; 6 Diagonal;
Tech	1 Repeat; 4 Gradient; 7 Texture;	2 Approximate; 5 Contrast; 8 Space;	3 Specific; 6 Emission; 9 Others
Contentproportion	1 below 25%; 4 above 75%	2 25-50%;	3 50-75%;
Graphics	1 Concrete;	2 Semi-abstract;	3 Abstract;
Colors	1 Warm;	2 Cold;	3 Neutral

TABLE 2 | Factor analysis.

Adjective	Factor 1	Factor 2	Factor 3
Traditional-Modern	0.788	0.041	0.015
Plain-Fashionable	0.744	0.241	0.081
Popular-Individualized	0.730	−0.015	0.292
Cold-Enthusiastic	0.100	0.871	0.055
Serious-Lively	0.089	0.868	0.087
Elegant-Heroic	−0.009	0.214	0.858
Soft-Intense	0.332	−0.063	0.775

shown in **Table 2**. The seven adjective pairs can be divided into three groups (KMO = 0.671, $p < 0.001$), and the interpretation rate of the total variance is 69.939%.

The evaluation criteria of the word pairs Traditional-Modern, Plain-Fashionable, and Popular-Individualized in Factor 1 is related to the social trend, so Factor 1 is named Sense of Contemporary. Factor 2 including Cold-Enthusiastic, Serious-Lively mainly describes the temperature of the poster, so Factor 2 is named Sense of Humanity. Factor 3 contains Elegant-Heroic and Soft-Intense that focus on the impact of the poster, so it is named Sense of Power. Take Factor 1, Factor 2, and Factor 3 as x, y, z coordinates, respectively, and the average of all samples in the three factors as the origin, the poster samples' visual imagery maps of every two aspects are drawn according to the difference (**Figure 2**). **Figure 2A** is the visual imagery map of ten poster samples at the technical aspect and the meaning aspect with the Sense of Contemporary and the Sense of Humanity as the dimensions. **Figure 2B** is the visual imagery map in the dimensions of Sense of Contemporary and Sense of Power. **Figure 2C** is the visual imagery map in the dimensions of Sense of Humanity and Sense of Power. According to **Figure 2**, the sense of contemporary of samples 8 and 9 increase. The sense of humanity of samples 1, 5, and 8 increased. The sense of power of samples 2, 4, 6, and 8 increased. The biggest change occurred in sample 8 in three dimensions. In general, all samples show a trend of moving closer to the center in the meaning aspect.

The independent-samples *T* Test of SPSS (IBM) was used to explore the differences between gender and profession ($n = 40$). It showed that there is no significant difference in the questionnaire results between different genders. Only in the meaning aspect, the professional group compared with the non-professional group consider that sample 5 was more fashionable ($F = 3.252$, $p < 0.05$). Therefore, this research will ignore the influence of gender and specialty on visual imagery. Using the paired-samples *T* Test of SPSS (IBM), the results show that the meaning aspect makes sample 1 more dynamic and enthusiastic, sample 2 more modern and intense, sample 5 more plump, sample 6 more personalized and fashionable, and sample 8 more dynamic. The visual imagery of the other samples doesn't change significantly.

DISCUSSION

The visual imagery at technical and meaning aspects of the poster is the emotional tendencies created by different colors, graphics, text, compositions, and themes. Current research shows that emphasizing emotion in design can help enhance the audience's acceptance and understanding of design works (Kao and Du, 2020). The experiment results show that audiences have different emotional characteristics for different presentation methods and themes. In terms of the technical aspect, it is clear that the visual imagery of the poster samples with rich colors and graphics are intense and lively, while that of the posters with poor elements are soft and serious (Ji et al., 2020). As for the meaning aspect, deep thinking enhances the sense of humanity and the power of posters related to social issues such as housing tension, tolerance, and respect for women, which means the posters became colder and softer for the audience in a way that posters related to water themes do not. An interesting thing to be concerned about is that different expressions will lead to different changes in visual imagery. Both samples 2 and 8 are calling on the protection of water. Sample 1 selects the cartoon image of a painful fish, and sample 8 selects the neat words of "WAR" and the graphics of a tombstone. The meaning of the posters increases the sense of power of sample 8 but decreases the sense of power of sample 1.

In the detailed comparative analysis of the technical aspect and meaning aspect, there is no significant difference in the visual imagery of samples 4, 7, 9, and 10. Considering the higher understanding degree of samples 4 and 7 than that of samples 9 and 10, there may be two reasons to explain this situation. One reason is that the audience did not grasp the deep meaning of the poster sample so the meaning aspect does not affect the final visual imagery. The second reason is that the visual imagery of the meaning aspect and technical aspect is consistent, so the emotion would not change. The comparison between samples 7 and 9 can verify the two reasons. In sample 7, the IKEA banners are entwined to form a monster roaring which covers the full poster, and the graphics are abstract with blue and yellow color. The audience's first feeling about the poster elements is conflict and

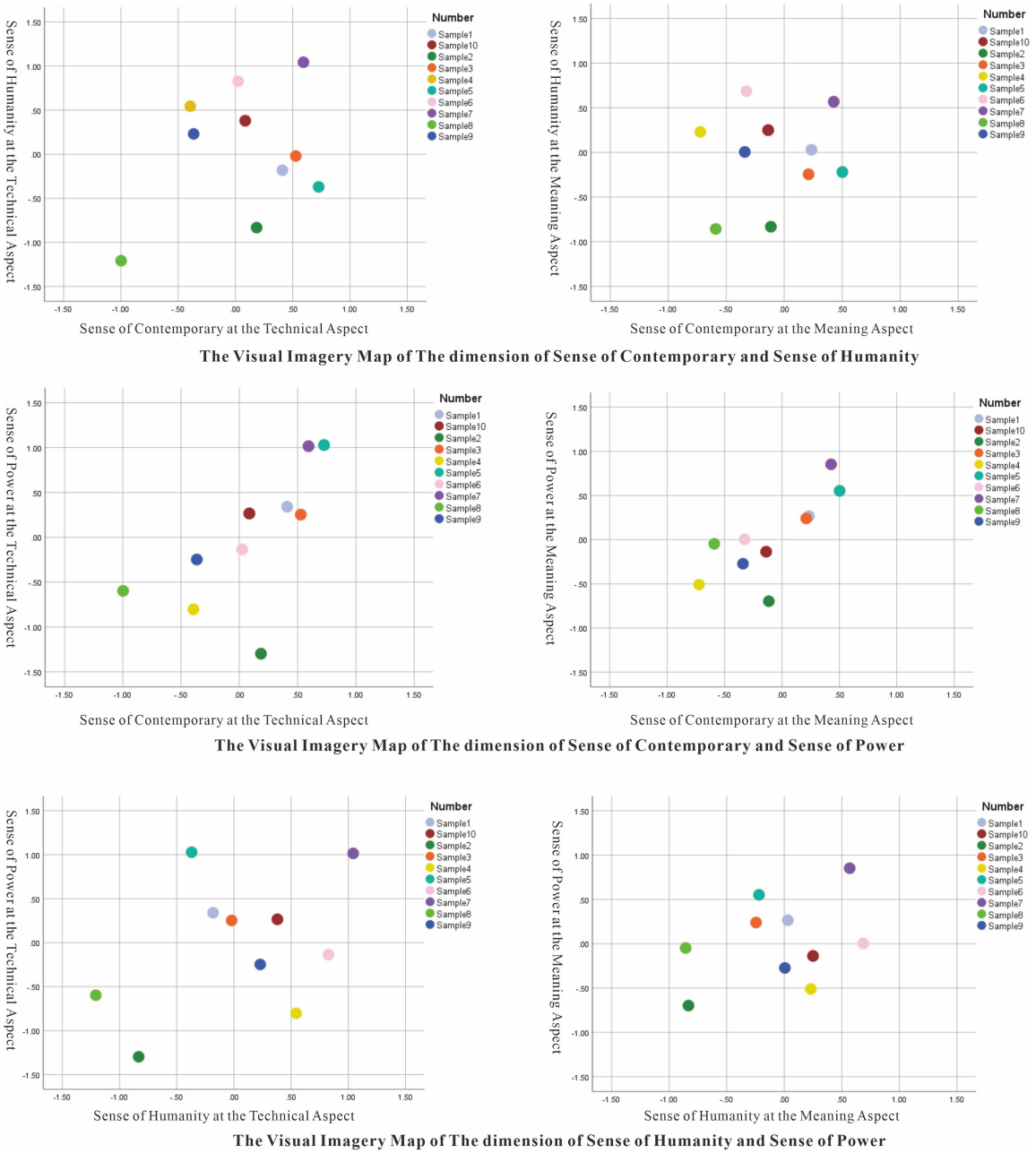


FIGURE 2 | Visual imagery map of the technical aspect and the meaning aspect (Poster source: https://www.anfachenaward.de/en).

strong which is consistent with the poster's meaning of housing tension and urban crowding. Even if some audiences do not notice the hidden IKEA logo and the vague monster image in the picture, they can still accurately receive the expected message of the poster. In contrast, sample 9 also expresses the theme Housing and Living where the text "We are all migrants" assists the semi-abstract graphics like a human-shaped airplane to express the housing crisis and housing instability. The visual imagery of the technical aspect with blue, white, and black color and the central composition conflicts with the strong visual imagery of the meaning layer about human existence. Even with the help of text, the understanding of the meaning of sample 9 is not better than that of sample 7.

It can be seen that only when the audience receives the information correctly, the meaning aspect affects the visual imagery of the technical aspect. And the poster with unified visual imagery at both aspects has a higher understanding. The consistency between the first impression and deep thinking, from the technical aspect to the meaning aspect plays a vital role in the persuasive effect for poster design.

CONCLUSION

This research found that different visual imagery will be produced by different colors, compositions, and themes of the poster. The audience will have a higher understanding of the poster that owns consistent visual imagery in the viewing process. The result suggests that the posters with the theme of women and children will awaken the softness of the audience, which can be assisted by softer colors and stable composition. The strong and conflicting visual effect will be more conducive to the expression of social themes such as environmental protection and social competition.

In conclusion, the same or similar visual imagery at the technical aspect and the meaning aspect encourage audiences to explore and think further. The posters that keep the visual imagery coherent can achieve a better effect.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author/s.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Ethics Review Committee of South China University of Technology. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

LW contributed to the conception of the study. WJ performed the experiment and the data analyses and wrote the manuscript. WC contributed significantly to manuscript preparation. SL helped perform the analysis with constructive discussions. All authors contributed to the article and approved the submitted version.

FUNDING

Foundation: Program of National Natural Science Foundation of China, No. 41971184; Team Project of Natural Science Foundation of Guangdong Province, No. 2018B030312004.

ACKNOWLEDGMENTS

The researchers would like to thank the 40 participants who participated in this experiment for providing vital data for this article.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2022.861366/full#supplementary-material

REFERENCES

Brown, K. G., Reidy, J. G., Weighall, A. R., and Arden, M. A. (2013). Graphic imagery is not sufficient for increased attention to cigarette warnings: the role of text captions. *Addiction* 108, 820–825. doi: 10.1111/add.12008

Charest, E., Gagne, M.-H., and Goulet, J. (2019). Development and pretest of key visual imagery in a campaign for the prevention of child maltreatment. *Glob. Health Promot.* 26, 23–31. doi: 10.1177/1757975917716924

Chen, Y., Yang, J., Pan, Q., Vazirian, M., and Westland, S. (2020). A method for exploring word-colour associations. *Color Res. Appl.* 45, 85–94. doi: 10.1002/col.22434

Ji, Q., Zhao, Y., and Liu, L. (2020). *Analysis of Color Language in Poster Design*. New York: Springer, 250–262.

Kao, T.-F., and Du, Y.-Z. (2020). A study on the influence of green advertising design and environmental emotion on advertising effect. *J. Clean. Prod.* 242.

Kim, H., and Fesenmaier, D. R. (2008). Persuasive design of destination web sites: an analysis of first impression. *J. Travel Res.* 47, 3–13. doi: 10.1177/0047287507312405

Molek-Kozakowska, K. (2020). Graphic emotion: a critical rhetorical analysis of online children-related charity communication in Poland. *Crit. Discourse Stud.* 17, 72–90. doi: 10.1080/17405904.2019.1567362

Purnomo, M., and Hidayatuloh, S. (2020). *Development of Rule-Based Expert System For Conceptualisation Of Poster Design. IOP Conference Series: Materials Science and Engineering*. Bristol: IOP Publishing, 012019.

Sentell, E. (2016). Making memories: writing and designing more memorable documents. *technical communication* 63, 136–153.

Tao, L. (2018). On Image: "Image" in meaning and visual rhetoric analysis. *J. Bimonthly* 149, 1–9.

Wu, T.-Y., and Liu, Y. (2016). *Comparative Study on Visual Differences of Poster Designs Based on Design Psychology. 18th International Conference on Human-Computer Interaction (HCI International)*. West Lafayette: purdue university.

Yang, L. (2016). *Audiences' Discussion on Visual Comprehend of Image Connotation–Examples from winning poster of Warsaw International poster biennale*. Guangzhou: Master, South China University of Technology.

Yang, Y.-F., Hsiao, W.-P., and Lu, J.-H. (2012). Discourse of design culture code. *Sustain. cities Dev. Env.* 209-211, 191–195. doi: 10.4028/www.scientific.net/amm.209-211.191

Yu, C.-Y., and Ko, C.-H. (2017). Applying facereader to recognize consumer emotions in graphic styles. *Complex Syst. Eng. Dev.* 60, 104–109. doi: 10.1016/j.procir.2017.01.014

Yufu, Y. (2002). *Design and Culture: design, sign, communication*. Beijing: Asia Pacific Book Press.

Zhang, P.-Z. (2015). *Research on poster style classification based on cluster analysis*. Qingdao: Master, Qngdao University.

Zhencheng, W. (2005). Image function in visual thinking – an interpretation of rudolph arnheim's visual thinking theory. *Academic Forum* 2, 129–133.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Wen, Jingjing, Chen and Luyu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers | Frontiers in Communication

ORIGINAL RESEARCH
published: 17 June 2022
doi: 10.3389/fcomm.2022.903015

Multimodal Prominence Marking in Semi-Spontaneous YouTube Monologs: The Interaction of Intonation and Eyebrow Movements

Stephanie Berger and Margaret Zellers*

Department of Linguistics and Phonetics, Institute for Scandinavian Studies, Frisian Studies, and General Linguistics, Kiel University, Kiel, Germany

Both facial expressions like eyebrow movements and prosodic characteristics like pitch height and the position of the pitch accent relative to the prominent syllable play an important role in prominence marking, which in turn is used by YouTubers and other public speakers alike to engage their audience. We analyze the co-occurrence of eyebrow movements and pitch accents of differing prominence levels produced by a sample of English-speaking YouTubers, and the height and position of those pitch accents. We annotated pitch accents, and analyzed videos using OpenFace 2.0, using three different eyebrow movement measures: eyebrow landmarks, distance between eyebrow and eye landmarks, and Action Units (facial expressions). We found that pitch height and eyebrow movements are positively correlated for at least some of the measures for all speakers. We also found that it is more likely for a medial or late pitch accent to arise at the time of an Action Unit connected to the eyebrows than an early accent, while there was no effect of the prominence level. However, we did not find evidence for a close temporal coordination of pitch movement and eyebrow movement. Finally, we found speaker-individual differences in the use of eyebrow movements and pitch accents, but also some slight indications of an effect for geographical origin of the speaker.

Keywords: eyebrows, intonation, pitch accents, OpenFace, prominence, multimodal, YouTube, monolog

OPEN ACCESS

Edited by:
Anna Esposito, University of Campania 'Luigi Vanvitelli, Italy

Reviewed by:
Manon Lelandais, Université de Paris, France
Iris Carmela Hübscher, University of Zurich, Switzerland

****Correspondence:***
Stephanie Berger
sberger@isfas.uni-kiel.de

Specialty section:
This article was submitted to Language Sciences, a section of the journal Frontiers in Communication

Received: *23 March 2022*
Accepted: *20 May 2022*
Published: *17 June 2022*

Citation:
Berger S and Zellers M (2022) Multimodal Prominence Marking in Semi-Spontaneous YouTube Monologs: The Interaction of Intonation and Eyebrow Movements. Front. Commun. 7:903015. doi: 10.3389/fcomm.2022.903015

1. INTRODUCTION

In this work, we analyze the co-occurrence of eyebrow movements and pitch accents of differing prominence levels, as well as the height of the pitch accents produced by a sample of English-speaking YouTubers. YouTubers—like others speaking in a public setting—rely on a combination of audio-visual cues in order to keep their audience engaged. Audience engagement is crucial since making videos on YouTube is not only entertainment, but also the YouTubers' business. Over the years, a very specific speaking style has been identified by newspapers and popular media, especially in YouTube videos where the YouTuber is talking directly to the camera in a monolog without props and in a semi-spontaneous, unscripted, but not completely unprepared style. It has been described as being "bouncy", a specific voice used for "talking to the audience"—some even call it "intellectual used-car-salesman voice" (Beck, 2015, no page numbers). These comments suggest that this particular speaking style is also quite controversial—it can be perceived as annoying

by many listeners/viewers (Jennings, 2021). This style of speech on YouTube has been dubbed "YouTube Voice". The label was first applied by Beck (2015), and then re-used by several other (online) news and lifestyle outlets (e.g., Green, 2015; Hagi, 2017; Jennings, 2021).

YouTube Voice is mainly characterized by a variety of emphasis strategies such as long and hyper-articulated vowels and consonants that YouTubers are "over-enunciating compared to casual speech which is something newscasters or radio personalities do" (Hagi, 2017, no page numbers; see also Beck, 2015). These hyper-articulated vowels and consonants are additionally used to create a certain rhythm that adds more interest to what the YouTuber is saying: regardless of the content, adding the rhythm and the "different kind of intonation makes it more engaging to listen to" and is meant "to show excitement and enthusiasm" (Hagi, 2017, no page numbers). Attracting and especially maintaining the attention of a viewer takes (vocal) effort, even when the viewer has clicked on the video of their own accord (Beck, 2015).

These emphasis strategies are of course not used exclusively on YouTube, but are standard strategies used for emphasis in everyday life. Emphatic accents like lengthened vowels (also termed *positive intensification*), lengthened onset consonants (*reinforcement accents*) or a number of these accents in a row creating a very noticeable rhythm (*accent chains*) appear everywhere for focus and to highlight important sections of an utterance (Kohler and Niebuhr, 2007; Niebuhr, 2010; Niebuhr et al., 2016; Berger et al., 2020). Studies have shown, for example, that the frequent use of emphatic accents like the ones mentioned above have a positive effect on a speaker's perceived vocal charisma, especially when the types are varied throughout (Niebuhr et al., 2016, 2020; Berger et al., 2020).

It is not only segmental emphatic accents that contribute to the perception of prominence. Pitch accents are local pitch maxima and minima that highlight words or syllables that are—mostly—of great importance for the information structure of an utterance. These pitch maxima and minima are associated with a syllable that is perceived as particularly prominent by listeners (Ladd, 2008; Sridhar et al., 2008). Additionally, the words or syllables that receive a pitch accent cannot always be predicted (Bolinger, 1972; Sridhar et al., 2008); their unpredictability can serve to draw more attention to a pitch-accented portion of an utterance. However, research likewise suggests that if a listener can entrain to the intonation contour a speaker uses *before* a prominent syllable, they "can predict where focus will occur in the utterance and get a head-start in navigating the discourse structure earlier on" (Ip and Cutler, 2021 p. 21, on Mandarin Chinese; see also Akker and Cutler, 2003, for similar results on English and Dutch). In their study, the accuracy of hearing a target phoneme correctly was higher in "accented contexts (10 misses)" than "in unaccented contexts (24 misses)" (Ip and Cutler, 2021, p. 15), suggesting that prediction accuracy is not 100%. In general—when nothing else changes in the contour—the higher a pitch accent is produced, the more prominent it is perceived (Gussenhoven, 2002). At the same time, pitch height goes in hand with pitch excursion.

The position of the pitch peak relative to the prominent syllable is also important, and, in some instances, can increase or even substitute the effect of pitch height on perceived prominence (Gussenhoven, 2002). If a pitch peak is higher, it tends to take the speaker longer to get to that higher point. It follows that higher peaks also tend to occur later in a prominent syllable than lower peaks. The longer rise takes more effort from the speaker which in turn causes late peaks to be perceived as more prominent than peaks that come early in or even before the prominent syllable (Gussenhoven, 2002). Because of the heightened production effort, "both higher and later peaks elicit more 'unusual occurrence' interpretations than 'everyday occurrence' interpretations" (Gussenhoven, 2002, p. 6). Additionally, pitch peaks marking new information in British English also tend to be produced later than other peaks, especially when it is the first pitch accent of an intonation phrase (Wichmann et al., 1997; Gussenhoven, 2002; Zellers et al., 2009). A recent study including the speakers of the present investigation found that the YouTubers in the sample indeed use significantly more late peaks than early peaks (pitch accents aligned with the prominent syllable being the most frequent peak placement), and likewise that emphatic (extra-strong) pitch accents were produced with higher pitch than accents with weak and strong prominences (Berger and Zellers, 2021). This terminology follows DIMA (*Deutsche Intonation: Modellierung und Annotation*, Kügler et al., 2019), see Section 3.2 for an overview of the annotation system.

Intensity has also been identified as a perceptual cue for prominence (e.g., Vainio and Järvikivi, 2006; Mo, 2008). The videos in this sample have different microphone set-ups and room acoustics, and would therefore not be comparable. Thus, the role of intensity in prominence is not discussed in detail here.

YouTube has the advantage of not only featuring the voice of the speaker, but also his or her gestures and facial expressions. We investigate eyebrow movements as facial expressions since they are always on screen in our data, and hand gestures frequently were moved outside the video frame. Eyebrow movements can be understood as co-called beat gestures that "do not necessarily convey any semantic content but are rather used to signal prominence, to construct rhythmical structures and to highlight words or expressions" (Ambrazaitis and House, 2017, p. 100; cf. McNeill, 2008; Loehr, 2012, for more details on beat gestures). There is much research that suggests a close (temporal) alignment of eyebrow movements and pitch accents, which is why eyebrow movements have often been seen as contributing to prominence, both in terms of the production as well as the perception of a speaker's prosody (e.g., Beskow et al., 2006; Flecha-García, 2010; Ambrazaitis and House, 2017).

Many studies have reported on the temporal alignment of eyebrow movements and pitch accents. One study found that in British face-to-face dialogues, the "start of the brow raise preceded very closely the start of the accented syllable" by about 60 ms, and that eyebrow raising was more frequent "when giving instructions than when asking a question or acknowledging the receipt of information" (Flecha-García, 2010, p. 551). That would suggest that eyebrows are used more when asking others to follow or engage. If that is the case, it is reasonable to assume

that eyebrow movements are also frequent on YouTube where the YouTuber asks the viewer to follow, continue watching, etc. Furthermore, Flecha-García (2010) also says eyebrow raises and prominence may be connected, perhaps as an extra level of prominence for especially important words or passages in addition to pitch accents. A study by Swerts and Krahmer (2010) further suggests that "the majority of the 'no accent' and 'weak accent' cases [in their data] occur without an accompanying eyebrow movement (76.8 and 62.8%, respectively), the strong accents more often co-occur with eyebrow movements (70.1%) than not. However, conversely, the mere presence of an eyebrow movement does not imply the presence of a strong accent, given that only a minority of the eyebrow movements (47 out of 303) occurs with a strong accent" (Swerts and Krahmer, 2010, p. 200f.).

While, hand gestures have been shown to align with (or adapt their alignment with regard to) prosodic phrasing (e.g., Loehr, 2004; Esteve-Gibert and Prieto, 2013; Krivokapic et al., 2015), cf. broader review in Wagner et al. (2014), there is little evidence supporting the association of eyebrow movements with prosodic phrasing. In Israeli Sign Language, eyebrow movements are a key component of intonational phrasing (Nespor and Sandler, 1999). However, since spoken and signed languages make different use of the bodily resources at hand, this cannot be automatically translated to spoken languages. Granström et al. (1999) report only a very weak correlation between eyebrow movements and prosodic phrasing in their data. Swerts and Krahmer (2008) report that the upper portion of the face, including the eyebrow area, is particularly important for prominence perception compared to the rest of the face, suggesting a possible asymmetry for how gesture types relate to prosodic structure.

Additionally, the frequency of eyebrow movements seems to differ between cultures: Swerts and Krahmer (2010) found a substantial amount of eyebrow movements in Dutch news readings, Flecha-García (2010) also found a substantial amount for English lectures and dialogues. Ambrazaitis and House (2017), on the other hand, found only very few eyebrow movements in their Swedish news data. However, the comparison to Swerts and Krahmer (2010) is tentative as the annotation methods differed between the two studies, even though the data were similar. In light of their results, Ambrazaitis and House (2017, p. 111) suggest that, in their data, eyebrow movements may "not at all have a prominence function proper [...], but rather represent an intensification signal which is added to a prominence" created by a pitch accent and head movement. Cavé et al. (1996, p. 2175) found that "rising-falling eyebrow movements produced by [their] subjects as they spoke were associated with F0 rises in only 71% of the cases. This suggests that eyebrow movements and fundamental frequency changes are not automatically linked [...], but are more a consequence of linguistic and communicational choices".

In the present study, we investigate eyebrow movements, their presence at the time of pitch accents, and differences in accent type and prominence level for a sample of six English-speaking YouTubers. Unlike the previous studies presented so far, we are not working with manually annotated gestures, but we investigate if automatically tracked facial landmarks and expressions are correlated with pitch accents. We therefore analyze videos with OpenFace 2.0 (Baltrusaitis et al., 2018), and focus on ten facial landmarks that are tracked by the software on the eyebrows, six distance measures between eyebrow landmarks and landmarks on the upper eye, and four Action Units (AUs)—specific facial expressions first introduced by Ekman and Friesen (1978) pertaining to the eyebrows, see Section 3.1 below for more detail.

In this study we investigate several research questions:

1. Is there a correlation between the movement of the eyebrow (operationalized as the height of the eyebrow landmarks, the distance between eyebrow and eye landmarks, and the intensity of the AUs) and pitch height at any given moment in the analyses?
2. What is the probability of an AU connected to the eyebrows co-occurring with accents depending on their prominence or accent type?
3. If there is eyebrow movement in the vicinity of a pitch accent, how close is the temporal alignment between them?

We hypothesize that there will be correlations between pitch height and eyebrow movement, and that the correlations suggest that eyebrows move up with higher pitch peaks. Furthermore, we hypothesize that pitch accents with a higher prominence level and appearing later relative to the prominent syllable likely occur at the same time as an Action Unit. Both of these hypotheses are based on the importance of pitch height and gestures for expressive speech, necessary for YouTubers and other public speakers. Additionally, we hypothesize that we will find peaks in eyebrow movement in the vicinity of pitch peaks, most likely preceding them, such that the eyebrow contour would peak and then be in a fall (or already complete) by the time of the F0 peak. We also hypothesize that this pattern might be stronger in higher-prominence pitch accents, and that different speakers might show different patterns of F0-eyebrow coordination. Finally, there might also be origin-, gender- or speaker-specific differences in the way pitch accents, prominences and eyebrow movements are used, which we will explore for our dataset.

2. MATERIALS

The data sample for this study consists of approximately 74 min of video material from six YouTubers. The videos came from YouTube in m4a format and were converted into wav files using FormatFactory (FreeTime, 2021) A subset of this sample is annotated for pitch accent placement and prominence levels (cf. Section 3.2) so that results on pitch accents are based on 3 min of annotated speech material (18 min total). Three of the YouTubers in the sample are from North America (two female—one from the US, one from Canada—, one male) and three are from England (one female, two male). All six of the YouTubers have business ventures outside of YouTube (such as books, tours, apparel companies, production companies, etc.). They have all been active on the platform for more than 10 years and grown their channels over time. Therefore, they all have considerable experience with public speaking (both online and in person at conventions), though it is likely they are not trained public speakers. **Table 1** summarizes the YouTubers' abbreviations used

TABLE 1 | Overview of YouTubers in sample, including identifying abbreviations, gender (f, female; m, male), origin (NAM, North America; ENG, England), channel with the analyzed video, and subscriber count of that channel (in million, obtained February 8, 2022).

Abbr.	Gender	Origin	Channel	Subscribers
CB	f	NAM	Colleen Ballinger	8.8M
DH	m	ENG	Daniel Howell	6.24M
LP	f	ENG	Louise Pentland	2.25M
LS	f	NAM	Lilly Singh Vlogs	2.77M
MF	m	NAM	Markiplier	31.9M
PL	m	ENG	AmazingPhil	3.98M

in the text, gender, origin, channels, and current subscriber count (date of subscriber count collection: February 8, 2022). The YouTubers were selected randomly to reflect a range of popularity with at least over 2 million subscribers. The videos were also selected randomly from the channels. The absence of sound effects and background noise in the majority of each video was priority. A range of topics is included, but all have emotional significance to varying degrees for the speaker.

All of the videos can loosely be categorized as vlogs (video blogs), where the YouTuber speaks directly into the camera more or less spontaneously, and talks to the audience member on the other side of the screen about a topic that is important to the YouTuber and their followers. These types of videos usually focus on interacting with the audience and calling the audience to join the community by subscribing or interacting with each other in the comment section of the video. All videos can be found on YouTube on the respective channels (see Ballinger, 2017; Howell, 2017; Pentland, 2017; Singh, 2017; Fischbach, 2018; Lester, 2018), but the videos are also provided in a playlist listed below.

3. ANALYSIS I: EYEBROW POSITION AND PITCH ACCENTS

3.1. Eyebrow Movement Analysis—OpenFace 2.0

In this investigation, eyebrow movements are tracked using OpenFace 2.0 (Baltrusaitis et al., 2018). OpenFace is an open access toolkit for the analysis of "*facial landmark location, head pose, eye gaze, and facial expressions*" (Baltrusaitis et al., 2018, p. 59, italics in original). Relevant for this study are facial landmark locations and facial expressions (Action Units). OpenFace's typical analysis uses a frame rate of 30 Hz (Cannata et al., 2020), but the outputs of the current analysis were not so consistent, ranging from 25 frames per second to 60 frames per second, which is likely caused by the inherent frame rate of the videos.

OpenFace tracks landmarks on the face, taking into account head pose as well. There are 5 landmark measurements for each of the eyebrows (in pixels) with x- and y-coordinates, of which we are only including the y-coordinates for vertical movements. Landmark y17 refers to the y coordinate of the left-most point of the eyebrow that is on the left of the video frame, while landmark y26 corresponds to the y coordinate of the right-most point of the eyebrow on the right of the video frame. This is shown in **Figure 1**, which is a representation of all facial landmarks recorded by OpenFace, as an example for the first analyzed frame of speaker PL. The landmarks referring to the eyebrows are marked in red. In order to gather more information on eyebrow movement, the distances between six of the eyebrow landmarks and landmarks on the eyes are also included following Nasir et al. (2016). In this way, we can also confirm that the movements of the landmarks are mostly connected to eyebrow movement, and not epiphenomena resulting from head movements. The distances are marked in green in **Figure 1**.

OpenFace also tracks so-called Action Units (AUs), based on the Facial Action Coding System by Ekman and Friesen (1978). AUs are stretches of time where some type of action happens on the face as a gesture. We analyze four different AUs connected to the eyebrows: inner brow raiser (AU01), outer brow raiser (AU02), brow lowerer (AU04), and upper lid raiser (AU05)—the latter is at the same time likely also associated with a widening of the eyes. OpenFace creates frame-by-frame pictures of aligned faces which feature only the face in the same position regardless of headpose, sized 112 by 112 pixels (Baltrusaitis et al., 2018, p. 63). OpenFace furthermore returns two variables for each of the AUs: whether or not they are present or absent in every analyzed frame, and the intensity on a scale of 0–5 which the action occurs with. Previous studies found that OpenFace performed AU detection at higher-than-chance levels, for posed datasets as well as videos collected "in the wild" from YouTube (Namba et al., 2021a). However, at least for still images, the accuracy drops when the face is angled at 45°, but still higher than chance (Namba et al., 2021b). Examples of the four Action Units are provided in **Figure 2**. The examples are taken from the aligned faces of three of the analyzed videos (Howell, 2017; Singh, 2017; Fischbach, 2018) and are examples of the specific AUs with the highest intensity when only one of the four AUs was detected, and not a combination of them.

We extracted the default data of an OpenFace analysis, but we are focusing our analysis on only a subset of the extracted features: the y-coordinates of the 2D facial landmarks 17–26 (eyebrows); additionally the y-coordinates of the upper lid landmarks 36, 37, 39, 42, 44, and 45; as well as the presence and intensity measures of Action Units 1, 2, 4, and 5.

3.2. Pitch Accent Annotation With DIMA

In order to investigate the relationship between prosodic structure, prominence, and eyebrow movements, pitch accents were annotated using the German consensus system DIMA (*German intonation: modeling and annotation*, Kügler et al., 2015, 2019; Kügler and Baumann, 2019). While, this system was created as an intonation annotation system for German, it is inherently based on a phonetic annotation: labels are placed when pitch movement can be heard by the trained annotators. In a second step, they are aligned and adjusted with the help of the visual pitch contour in Praat (Boersma and Weenink, 2018). Unlike other systems for intonation annotation like ToBI, there are no phonological decisions made at the time of the annotation, but they are possible in a second step (see Kügler et al., 2019 for a

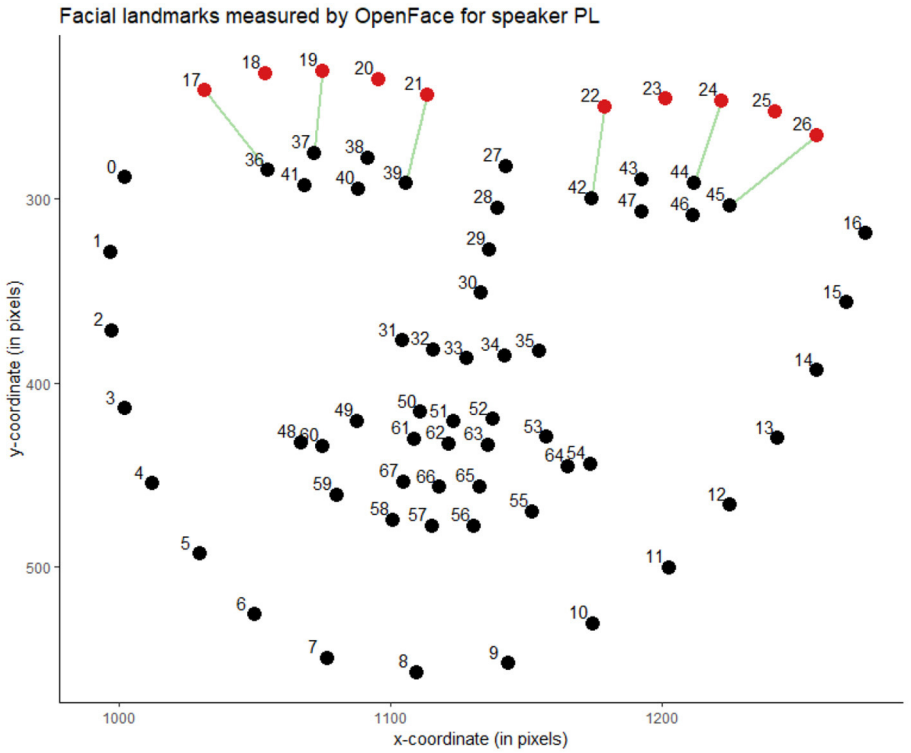


FIGURE 1 | The facial landmarks measured by OpenFace, exemplified for the first frame of the video featuring speaker PL (Lester, 2018). The eyebrow-specific landmarks are highlighted in red, and the numbers used in the analyses to refer to the landmark coordinates are indicated. The green lines represent distances measured between two facial landmarks.

FIGURE 2 | Examples of the four Action Units (AUs) investigated in this study, taken from the aligned faces extracted by OpenFace: **(A)** AU01 (inner brow raiser, measured intensity: 2.51); **(B)** AU02 (outer brow raiser, measured intensity: 1.06); **(C)** AU04 (brow lowerer, measured intensity: 3.46); **(D)** AU05 (upper lid raiser, measured intensity: 3.07).

comparison of DIMA with other intonation annotation systems). Therefore, the system can be easily applied to other languages such as in this case, English.

The following description of the annotation process is a general summary of the elements of DIMA. All information hereafter is based on the annotation guidelines—in German—by Kügler and Baumann (2019). Only the symbol inventory relevant for this study is included. For all other elements of the system, please consult the guidelines or the overview of the system's symbol inventory in Kügler et al. (2019).

DIMA works with a minimum of six annotation tiers. Two of these tiers—a word-level and a syllable-level segmentation of the speech material—should be available before the start of the intonation annotation. The DIMA-specific tiers are a phrase level, a tone level, a prominence level, and a last tier for comments where uncertainties and observations can be noted. In a first step, all labels are placed purely on an auditory basis, and are then adjusted and aligned with the visible pitch contour. The labels are placed in a specific order of tiers. First, minor and major phrases are delimited by boundaries on the phrase tier (based on pauses, phrase-final lengthening, F0 reset, and the strength of the audible break). Next, prominent syllables are identified, and their strength is annotated; the strength of a prominence is referred to in the rest of the work as prominence level. Non-prominent syllables are not labeled at all. Weakly prominent syllables are labeled "1". Syllables with a strong prominence get the label "2". Usually, prominences of level 2 are associated with an accent tone, though there can be exceptions. Extra-strong or emphatic prominences are labeled "3". These prominences are labeled when an accent tone occurs and additionally, the prominence is increased by other factors like, for example, extreme pitch height, extreme loudness, segmental lengthening

of the consonant onset or the vowel of the prominent syllable, etc. (cf. Kohler and Niebuhr, 2007; Niebuhr, 2010). Segmental lengthening was perceptually determined. Measurements were carried out in a next step.

Finally, after the prominences, the tones are labeled on the tone tier. Tones are labeled with a "H" for high tones and "L" for low tones, plus frequently occurring diacritics. There are three types of tones:

- Boundary tones (placed at the same time point as the boundary labels on the phrase tier),
- Accent tones (placed at pitch accents, F0 maximum, or minimum), and
- Non-accent tones (placed at turning points—maxima or minima—of the pitch contour between pitch accents).

The tones are modified by diacritics in reference to the last tone of the same quality (high or low) that occurred before, though these modifications usually occur with high tones, marking downstep and upstep ("!" and "^", respectively).

The current study only investigates high accent tones, including tones with an upstep or downstep. We do not make major functional distinctions here, although the degree of prominence might partly account for the contrast between broad and narrow focus. Accent tones are labeled with an asterisk (e.g., H*) to mark the association of the tone with the prominent syllable. Accent tones will also be referred to as pitch accents in the remainder of the study. The placement of the accent tone in reference to the prominent syllable is also annotated. The starred tone label is always placed in the prominent syllable. If the F0 maximum occurs in the prominent syllable, the label is placed on the maximum. If the maximum occurs in the syllable before or after the prominent syllable, but is still associated with the prominent syllable, the H* label is placed in the center of the prominent syllable. A second label with a diacritic is then placed at the maximum in the syllable before (>) or after (<) the prominent syllable.

3.3. Data Treatment

All data treatment procedures were carried out in R Studio (RStudio Team, 2021). The raw OpenFace results for each video were reduced to the information relevant for the study: frame, timestamp, confidence, landmarks y17-26 for eyebrow height, and y36, y37, y39, y42, y44 and y45 for the distance measures, as well as AU01/02/04/05_r for the intensity of these Action Units, and AU01/02/04/05_c for their presence or absence. The timestamp column was also used to extract fundamental frequency (F0) measurements at the same timepoints that OpenFace measured. The F0 measurements were extracted by a Praat (Boersma and Weenink, 2018) script written by the authors. It extracted the F0 value at the timestamp both in Hertz (Hz) and in semitones re 100 (st).

The two datasets (OpenFace and F0 measurements) were processed and merged in R. The first round of data processing was run for each speaker individually. When the Action Unit intensity was smaller than 0.01, the value of the corresponding presence column was set to 0. When the intensity was bigger or equal to 0.01, the corresponding value in the presence column was set to 1. Duplicate rows were excluded by checking the frame number. Additionally, frames where the confidence of the OpenFace algorithm was below 0.85 were also excluded from the data (following Yunus et al., 2021). For two speakers (DH and PL), sections of the original video with scripted content, sound effects or background music were excluded from the dataset as well. The first 1.3 s of material from speaker LS were also excluded as OpenFace tracked the Smiley face on the speaker's shirt instead of her face. This was checked by visually inspecting the aligned faces images saved by OpenFace. Once this was finished for all speakers, the separate datasets were bound into one large dataset. The dataset was saved and the timestamps were manually adjusted with the original frame rate in a new column for all speakers so that no gaps with excluded frames appear in visualizations. Finally, columns with the distances between the eyebrow landmarks and the landmarks of the top eyelids were calculated by subtracting the eyebrow landmark location from the eye landmark location.

The pitch accents and their distance from the AUs were also measured. A Praat script written by the authors added the AUs as annotations into the Praat TextGrid that already contained the DIMA annotation. For each of the four AUs, two TextGrid tiers were added: one interval tier where the beginning and end of an Action Unit was annotated, labeled with the AU abbreviation, and one point tier where each measured timepoint was labeled with the intensity of the AU, only when an AU occurred. **Figure 3** shows an example annotation of the DIMA annotation and the added AUs. The tiers with the AU intensities at each timestamp are not included in **Figure 3** for ease of reading. Another script then measured for each pitch accent if there was an active AU at the time of the pitch accent and, if so, what the time distance was between the accent and the beginning/end of an AU. If there was no AU at the time of the accent, the time distance to the beginning of the following and the end of the preceding AU was measured. All time distance measurements were calculated twice in order to compare results for different labeling positions: once for the time of the pitch accents' association to the prominent syllable, and once for the time of the pitch peak (which could occur either within the prominent syllable, or the syllable before or after). The prominence levels of the pitch accents were measured in a separate step, as were the pitch values at the time of the pitch peak. The three resulting datasets were then merged for further data processing.

The same data that was removed in the OpenFace dataset was also removed for the PitchAccent data. For this study, as mentioned above, only high pitch accents were included in the analyses, and therefore a subset was created. The numerical prominence levels (1, 2, 3) were renamed into their connected categories "weak", "strong", and "emphatic"; the levels and the category names will be used interchangeably throughout the analysis. Values of time differences between an accent outside an AU and the beginning/end of the nearest AU were excluded if the time distance was more than 3 s away from beginning/end, following Flecha-García (2010). The remaining values were counted as occurring in the vicinity of an Action Unit. The PitchAccent data comprise around 180 s of speech material for

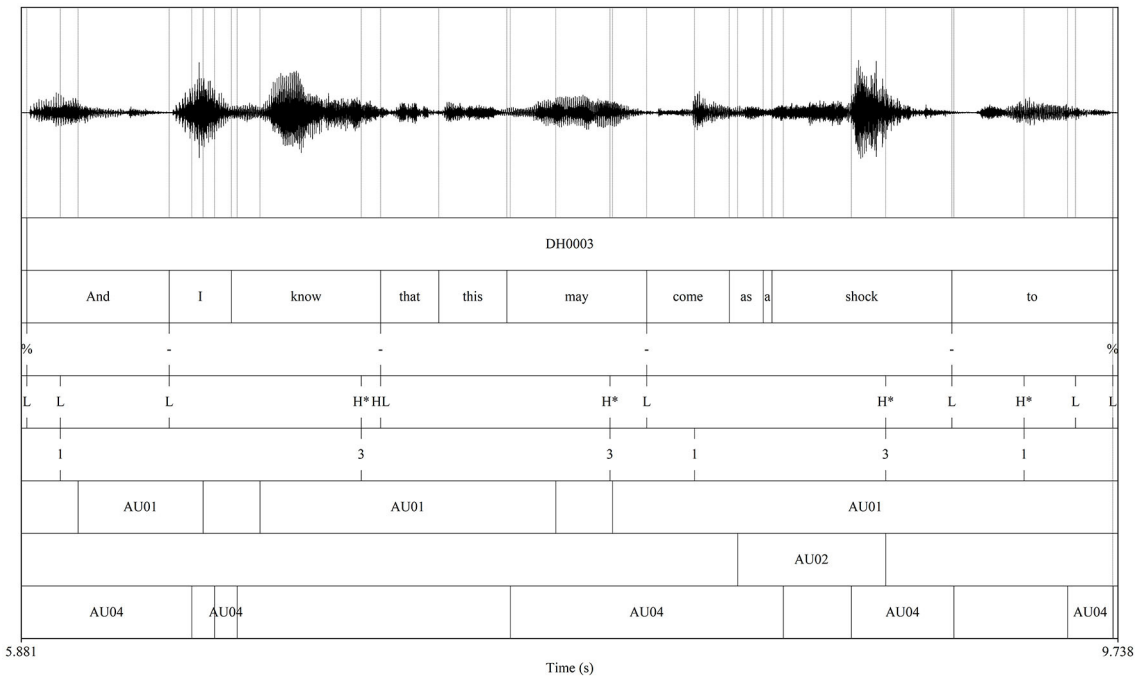


FIGURE 3 | An example annotation with DIMA and the AUs added by a script. Boundaries for the AU intervals were placed at the first and last points an AU occurred. For ease of readability, the AU05 tier is removed here as it was not present in the selection.

each of the six speakers, the amount that DIMA annotations are available for.

In total, 49 pitch accents were removed because pitch could not be measured. Additionally, 5 data points were lost while building the PitchAccent dataset. The exclusion of OpenFace data with a confidence below 0.85 caused the removal of 534 rows across all speakers, the majority with speakers DH and PL. **Table 2** shows the measuring points that were left in the OpenFace data after data processing, as well as the number of pitch accents and measuring points within the 3 min of annotated DIMA material per speaker together with the actual length of annotated material. Note that speaker MF has over twice the amount of measuring points as the other speakers. We assume that this comes from the internal frame rate used for his video, and that the other speakers used different frame rates. The OpenFace settings were the same for all six speakers.

TABLE 2 | An overview of the analyzed material including speaker abbreviation, the number of pitch accents and the length of DIMA-annotated material per speaker as well as the number of measuring points (= MP) of the OpenFace analysis both for the full video (after the exclusion of frames because of measuring issues or low confidence, see description above) and for the length of the annotated material.

Speaker	Pitch accents	Length material	OpenFace MP	OpenFace MP in annotated material
CB	242	185s	7,560	4,612
DH	202	184s	7,926	4,601
LP	216	181s	9,019	5,418
LS	248	181s	7,197	4,267
MF	255	184s	18,207	11,041
PL	242	180s	7,360	4,501

3.4. Results

Several statistical approaches are combined in this study which are reported in the corresponding sections below. Section 3.4.1 will first look at the correlations between pitch height and the intensity of Action Units, the eyebrow-specific facial landmarks, as well as the distances between eyebrow and eye landmarks. Then, the results from binomial regression models investigating the likelihood of a pitch accent occurring at the same time as an AU are investigated in Section 3.4.2.

3.4.1. Correlations of Pitch and Eyebrow Movement

We carried out Pearson correlations between pitch (in semitones) and the intensity of the Action Units, the eyebrow-specific facial landmarks as well as the distances between the eyebrow- and eye landmarks using the `cor.test()` function in R and visually corroborated them with scatter plots. The correlations were run on subsets of the data containing only one of the four Action Units or only the landmarks. Thus, zero values due to non-existence of the specific Action Unit could be excluded without losing data when other Action Units were present. Pitch measures that were undefined because pitch could not be measured (due to creaky voice or unvoiced segments) were excluded. Therefore, the correlations and their visualizations are investigating what the correlation is when there *is* measurable pitch and there *is* an Action Unit. The same is the case for the landmarks and landmark distances: we look at correlations of eyebrow

TABLE 3 | Correlation coefficients for the correlations between pitch (in semitones) with the Action Units (AU01–AU05) and the eyebrow-specific facial landmarks (y_17–y_26) for each of the six speakers. Additionally, the correlations between pitch and the distances between the eyebrow and the eye landmarks—used as a measure for eyebrow movement—are included.

	Full sample	CB	DH	LP	LS	MF	PL
AU01	−0.025	−0.005	−0.006	−0.021	0.086	0.149	0.017
AU02	−0.115	0.017	−0.007	0.058	0.186	0.148	−0.011
AU04	0.030	0.287	0.231	0.01	0.176	0.297	0.013
AU05	0.209	0.093	0.099	0.102	0.38	0.072	0.08
y_17	−0.29	−0.123	−0.002	−0.066	−0.26	−0.135	−0.071
y_18	−0.291	−0.12	−0.007	−0.069	−0.259	−0.132	−0.069
y_19	−0.287	−0.115	−0.011	−0.07	−0.261	−0.129	−0.067
y_20	−0.295	−0.107	−0.014	−0.072	−0.263	−0.132	−0.064
y_21	−0.304	−0.098	−0.017	−0.073	−0.265	−0.139	−0.062
y_22	−0.317	−0.07	−0.022	−0.076	−0.263	−0.136	−0.055
y_23	−0.326	−0.059	−0.025	−0.075	−0.257	−0.129	−0.051
y_24	−0.33	−0.045	−0.029	−0.075	−0.25	−0.118	−0.048
y_25	−0.333	−0.033	−0.034	−0.077	−0.244	−0.107	−0.048
y_26	−0.335	−0.025	−0.035	−0.077	−0.239	−0.106	−0.048
y_36-y_17	0.105	0.061	−0.038	0.010	−0.49	−0.007	0.051
y_37-y_19	0.091	0.030	0.013	0.008	0.034	0.048	0.058
y_39-y_21	0.131	0.019	−0.008	0.014	0.142	−0.028	0.021
y_42-y_22	0.117	0.33	−0.028	0.013	0.173	0.015	0.030
y_44-y_24	0.094	0.074	−0.019	−0.003	0.153	0.127	0.072
y_45-y_26	0.101	0.104	−0.040	−0.015	0.014	0.027	0.048

The shaded areas indicate significant correlations ($p < 0.05$).

movements when there *is* measurable pitch. This analysis was run for OpenFace and pitch data of the entire video durations, meaning about 74 min total (after exclusion of scripted material and analyses with lower than 0.85 confidence).

Table 3 summarizes all correlation coefficients for each of the speakers. Significant correlations are shaded ($p \leq 0.05$). All of the speakers had at least one significant correlation for pitch with an Action Unit, and at least three significant correlations for pitch with an eyebrow-specific facial landmark. The significant correlations between AUs and landmarks with pitch differ depending on the speaker, as does the strength of the correlation and also, for AU01 and AU02, the direction of the correlations.

AU05, the upper lid raiser, is the only Action Unit with significant correlations across all speakers. This correlation is positive for all speakers, suggesting that the higher the pitch is, the more intense is the raising of the eyebrows. **Figure 4** shows the correlations between pitch and AU05 for each of the speakers.

Additionally, the facial landmark for the middle brow marker of the rightmost brow in the video (y_24) and the one next to it to the right (y_25) are the only two facial landmarks that show significant correlations with pitch for all speakers. A *lower* pixel value on the y-axis corresponds to a *higher* placement of the landmark in the face (cf. **Figure 1**). The correlations of y_24 and y_25 are negative for all speakers, suggesting that the pixel value is lower the higher the pitch is. Keeping in mind the actual correspondence of pixel value to landmark placement, that actually means that the eyebrows are higher when the pitch is higher. **Figure 5** illustrates the correlations of landmark y_24 for each of the speakers.

Speakers MF and LS have significant correlations for pitch with all 14 other features, both AUs and facial landmarks (all $p \leq 0.001$). These two speakers also consistently have the strongest correlation coefficients, LS stronger than MF. Additionally, speakers LP and PL's correlations of pitch with facial landmark features are all significant (all $p \leq 0.002$). This is similar for speaker CB, with the exception of the correlation of pitch with the rightmost eyebrow landmark (y_26) which is not significant. For speaker DH, only the three facial landmarks to the right of the video frame (y_24–y_26) are significant (all $p \leq 0.05$).

The speakers from North America (MF, LS, and CB) have stronger correlations than the other three speakers from England. These three speakers from North America consistently have correlation coefficients stronger than 0.1. For the speakers from England, this is not the case. Speakers DH and LP each have one correlation of comparative strength (all other correlation coefficients are weaker than 0.1), speaker PL does not have any correlation coefficients above 0.1. There is no obvious pattern according to gender.

The distances between the eyebrow and eye landmarks overall correlate significantly with pitch in semitones (all $p < 0.001$). The correlations of most of the distances are positive, suggesting the higher the pitch, the larger the distance between the two facial landmarks. Only the distance between the rightmost eyebrow landmark (y_26) and the rightmost eye landmark (y_45) was negatively correlated, suggesting that as the pitch increased, the distance between the two landmarks got smaller.

While the correlations are significant for the sample as a whole, there are strong differences between speakers. For speaker LP, none of the correlations between distance measures and pitch were significant. For all other speakers, there were at least two significant correlations. Speaker DH has significant correlations for the two outermost eyebrow landmarks (y_17 and y_26) with the respective eye landmarks. Both of these correlations are negative. For the other four speakers, the significant correlations differ between landmarks. However, all four speakers have significant positive correlations of the center eyebrow landmarks (y_19 and y_24) with the respective eye landmarks (y_37 and y_44) in common, suggesting that for these speakers, the middle part of the eyebrows tends to be raised when the pitch is higher. Another observation is that the two speakers with few to no significant correlations are two of the three speakers from England in the sample, perhaps hinting at less facial expressivity (at least in terms of eyebrow movements) for English speakers compared to North American speakers.

Since we expect greater expressivity to be associated with more perceived charisma (see, for example, Niebuhr et al., 2020), a PASCAL score (*Prosodic Analysis of Speaker Charisma: Assessment and Learning*) was calculated based on 38 acoustic-prosodic parameters weighted by the PICSA algorithm (Niebuhr et al., 2017; Niebuhr, 2021). The measurements going into the analysis were mean values and absolute frequencies calculated

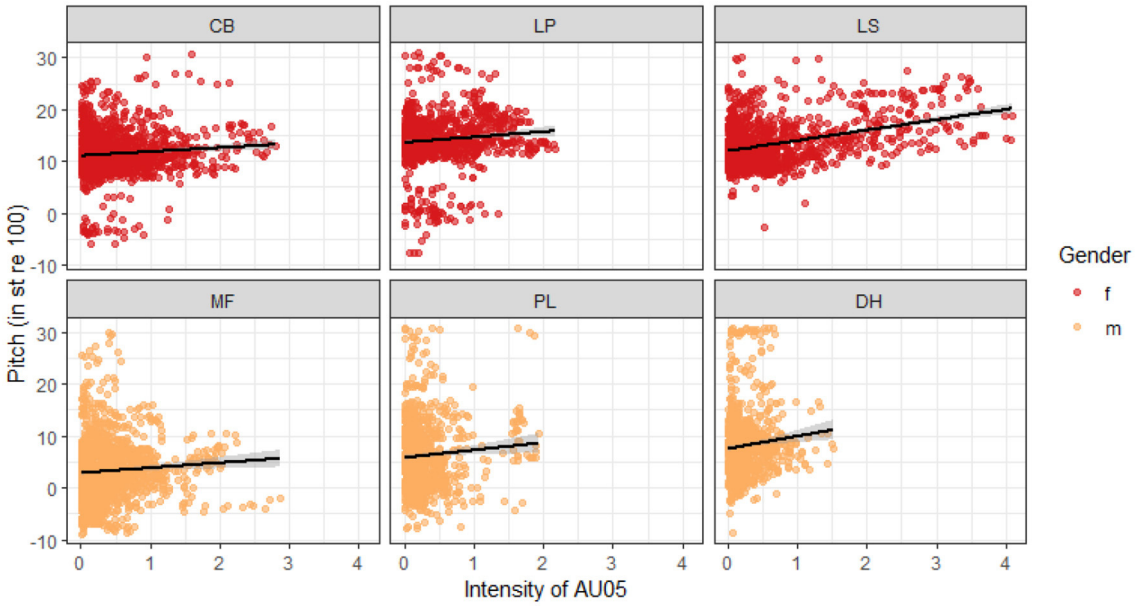


FIGURE 4 | Correlations between pitch (in semitones) and the intensity of AU05 (upper lid raiser) for each of the six speakers.

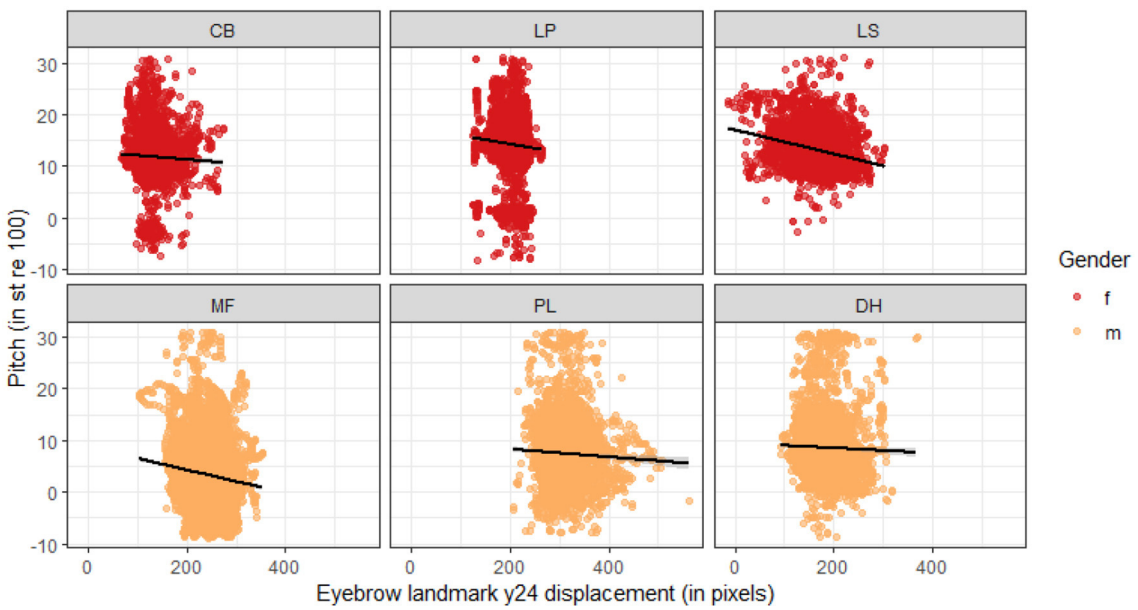


FIGURE 5 | Correlations between pitch (in semitones) and the displacement of the eyebrow landmark y_24 (middle landmark on rightmost eyebrow, in pixels) for each of the six speakers. Negative correlations correspond to a raising of the eyebrow landmark as pitch increases: The y-coordinates have lower values the higher the landmark is in the video frame, and thus the face.

for the first two of the 3 min of annotated speech material for each of the six speakers. The charisma score is a number between 0 and 100. All of the six speakers are close together in their scores (see **Table 4**), around the middle of the score scale. The speaker with the highest charisma score in the sample (DH) had the fewest correlations between pitch and AUs or facial landmarks. The speaker with the second-highest charisma score (MF) and the speaker with the lowest charisma score (LS) have the most correlations of all speakers, suggesting expressive pitch and expressive eyebrows.

TABLE 4 | PASCAL scores for the different speakers in the sample (out of 100).

Female speakers		Male speakers	
CB	57	DH	61
LP	55	MF	59
LS	45	PL	53

3.4.2. Likelihood of Pitch Accent Occurring During Action Unit

We also ran binomial regression models with the `glmer()` function (Bates et al., 2007) in R. This method tests if a pitch accent of a certain prominence level (weak, strong or emphatic) can predict the presence or absence of an Action Unit. The models were calculated both for the position of the pitch peak (within the prominent syllable, or in the syllable before or after the prominent syllable) as well as for the time of a prominent syllable each pitch peak was associated with. Additionally, we ran further binomial regression models testing if a pitch accent of a certain type (early, medial, late) can predict the presence or absence of an Action Unit. These models were only calculated for the time of the position of the pitch peak. All models contained the speaker as a random factor. The models performed best when compared to a model without random factor and to a null model (established by a smaller AIC value, e.g., for AU01 and accent type: $AIC_{null} = 1949.1$, $AIC_{glm} = 1948.3$, $AIC_{random} = 1911.6$). There was no significant difference between the models referring to the pitch peak or the prominent syllable, so only the results of the models referring to the position of the pitch peak will be reported. Examples 1 and 2 below show the R formulas used for the binomial regressions of AU05 with prominence level and accent type.

$$glmer(AU05.pres.peak \sim prom + (1|speaker), \\ data = pitchaccents, family = binomial) \quad (1)$$

$$glmer(AU05.pres.peak \sim acc.type + (1|speaker), \\ data = pitchaccents, family = binomial) \quad (2)$$

The binomial regression models revealed no significant effects of the prominence level. The presence of a pitch accent with a particular prominence level therefore does not predict the occurrence of an Action Unit (all $p \geq 0.1$). This can also be seen visually in **Figure 6**, where pitch accents seem to co-occur with AUs, but sporadically and with no obvious prominence pattern. Therefore, there seems to be no one-to-one correlation between the presence of a pitch accent of any prominence level and the presence of an Action Unit. However, we assume that there is a correlation with the accent type and an AU.

For accent type, the binomial regression models revealed significant effects for AU01 (inner brow raiser) and AU05 (upper lid raiser). There were no significant effects for the other two AUs. The model for AU01 suggests that the presence of a medial pitch accent (H^*) predicts the occurrence of an inner brow raiser ($z = 2.118$, $p = 0.03$). For AU05, the model suggests that not only does the presence of a medial peak predict the occurrence of an upper lid raiser ($z = -2.807$, $p = 0.005$), but that the presence of a late peak ($H^*<$) likewise predicts the occurrence of an upper lid raiser ($z = -4.398$, $p < 0.001$). Visually, this result is not obvious, though, at least not in the example (see **Figure 7**).

Additionally, the fact that the models with speaker as a random factor performed better than the models without suggests that the speaker is also a source of substantial variation which cannot be accounted for by the prominence level or accent type.

The logistic regression models only test the presence or absence of an AU at the time of the peak. To investigate the actual timing of pitch accent peaks and eyebrow movement peaks, the next analysis, in Section 4, looks at AUs in the vicinity of pitch accents, taking prominence into consideration.

4. ANALYSIS II: FUNCTIONAL DATA ANALYSIS

The degree to which different states (e.g., high F0 or raised eyebrows) occur at a specific point in time is not the only parameter of interest in an investigation of timing. The degree to which changes in these parameters may correlate with each other across time also plays an important role. For example, the claim that eyebrow movements are aligned with F0 movements (Flecha-García, 2010) suggests that a rise in F0 should be simultaneously accompanied, or perhaps preceded by at a relatively fixed time point, by a lifting of the eyebrows. In order to test such a hypothesis, methodologies are needed which can investigate the relationships between data in the form of contours rather than individual points. One such collection of methodologies is assembled under the name of Functional Data Analysis (FDA) (Ramsay and Silverman, 2005).

Functional Data Analysis is a set of methodologies allowing for the extension of classic statistical tools to the domain of functions. Entire contours—represented by functions—are the input to the analyses, rather than data points. Since the analysis is conducted automatically with minimal input from the analyst, this method is much less prone to subjective bias, and can also identify patterns which may not immediately arise from a visual inspection of the data.

FDA has been used to model variation in prosody across a variety of domains and languages, including the investigation of tone dynamics in Taiwan Mandarin (Cheng et al., 2010; Cheng and Gubian, 2011); the analysis of intonational categories in, for example, Italian (Turco et al., 2011) and Greek (Lohfink et al., 2019); and the analysis of F0 variation in the context of new-topic initiation in Southern British English (Zellers et al., 2010). Multidimensional analyses, involving more than one contour, are also possible. Thus FDA has also been used to investigate formant transition features. First and second formants are used together to distinguish contexts with deleted schwas from contexts where no schwa was present in French (Gubian et al., 2009) as well as to distinguish diphthongs from pairs of adjacent vowels in Peninsular Spanish (Gubian et al., 2015). A combination of the parameters F0 and speech rate was used to

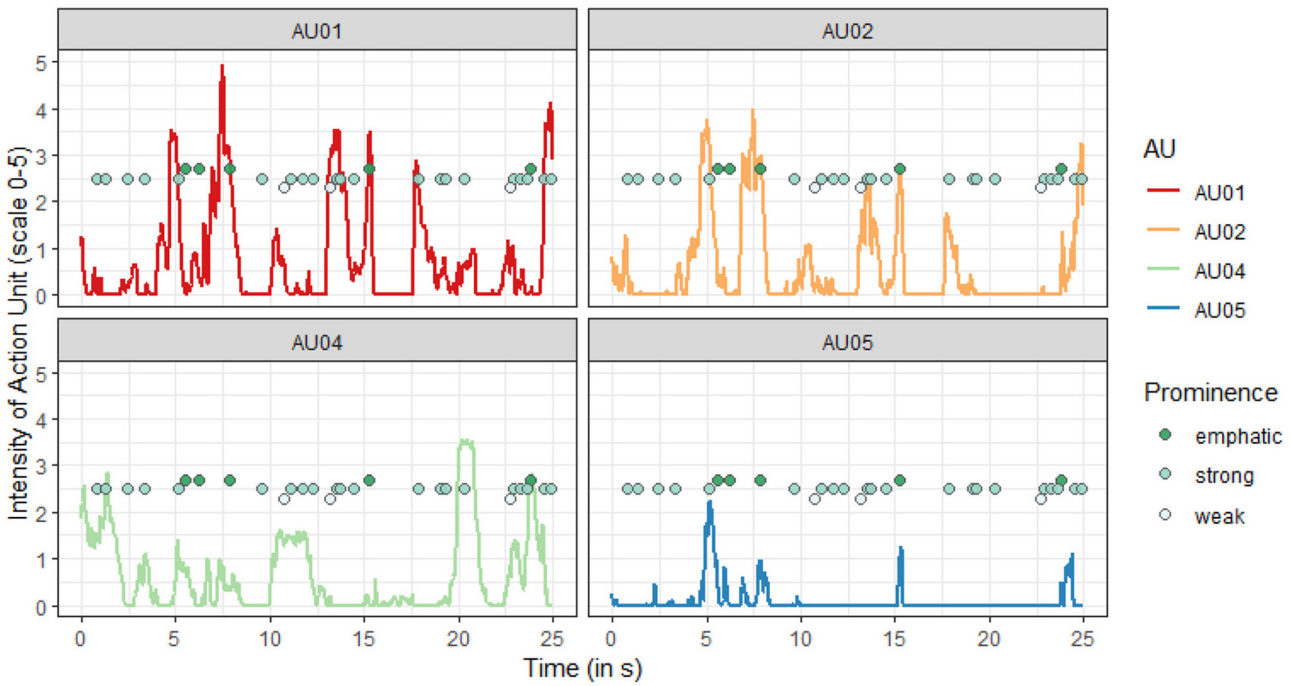


FIGURE 6 | Co-occurrence of Action Units and pitch accents with their prominence level (example). The intensity of each of the four Action Units (scale of 0–5) is shown for the first 25 s of speech material of speaker MF. The pitch accents and their prominence level are super-imposed, the y-axis does not refer to the prominence level directly, but emphatic accents are placed higher than accents with strong prominence, which in turn are placed higher than accents with weak prominence. The time of the pitch accents is that of the pitch peak, so either within the prominent syllable, or in the syllable before or after.

investigate how first language prosody influences production of hyperarticulation of prosody in a second language (Asano and Gubian, 2018).

We carried out an instrumental acoustic analysis of the pitch contours and the eyebrow movements using tools from Functional Data Analysis. The specific analysis carried out was a Functional Principal Component Analysis (FPCA). Principal component analysis takes a multidimensional dataset and uses orthogonal transformations to make sources of variation more accessible; this can be visualized as replotting the data on a new set of axes, which are the dimensions along which the most variation is present in the data. These new axes are the Principal Components (PCs). In FPCA, the PCs allow us to visualize aspects of contour shape which are mathematically relevant to the variation across a set of contours.

The FPCA was carried out in R (RStudio Team, 2021) using the scripts provided by Michele Gubian (available https://github.com/uasolo/FDA-DH/).

4.1. Data and Methodology

The input to the FPCA analysis comprised two datasets which were coordinated in time. The data were selected by locating the F0 peaks annotated for the first portion of the analysis (see Section 3.2). A time window of 1 s was identified with the F0 peak falling exactly in the middle (i.e., the window extended 500 ms to the left and to the right of the F0 peak). A 1-s window was chosen as a window that was large enough to potentially identify some systematicity in the forms of the contours while not overlapping too much with movement associated with adjacent pitch accents; since the mean distance between adjacent pitch accents varied between speakers from 0.61 to 0.83 s in the current data, a smaller window might have been even more ideal, but might have resulted in too much data loss to be able to carry out the analysis.

In order to align the video and the audio measurements, the time points used to extract the points comprising the contours were determined based on the output of the OpenFace analysis, which were 4 ms apart, resulting in a maximum of 25 time points per sample. The time points at which measurements of the eyebrow position were available were extracted from the OpenFace data. Using a Praat script, F0 measurements were taken at time points which fell within the windows identified around the F0 peaks; the measurements were taken in semitones re: 100 Hertz using Praat's automatic parameters. These F0 measurements were then re-combined with the eyebrow movement data in R, resulting in parallel simultaneous measurements for F0 and the four tested AUs at the times surrounding the F0 peaks. In this way, "contours" were created for both features, where the X-value was the time point, and the Y-value respectively either the F0 measurement or the value of the AU (between 0 and 5).

Since the stretches of speech were not controlled in terms of their segmental content, some of the extracted F0 values were missing due to voiceless segments or creaky voice. For

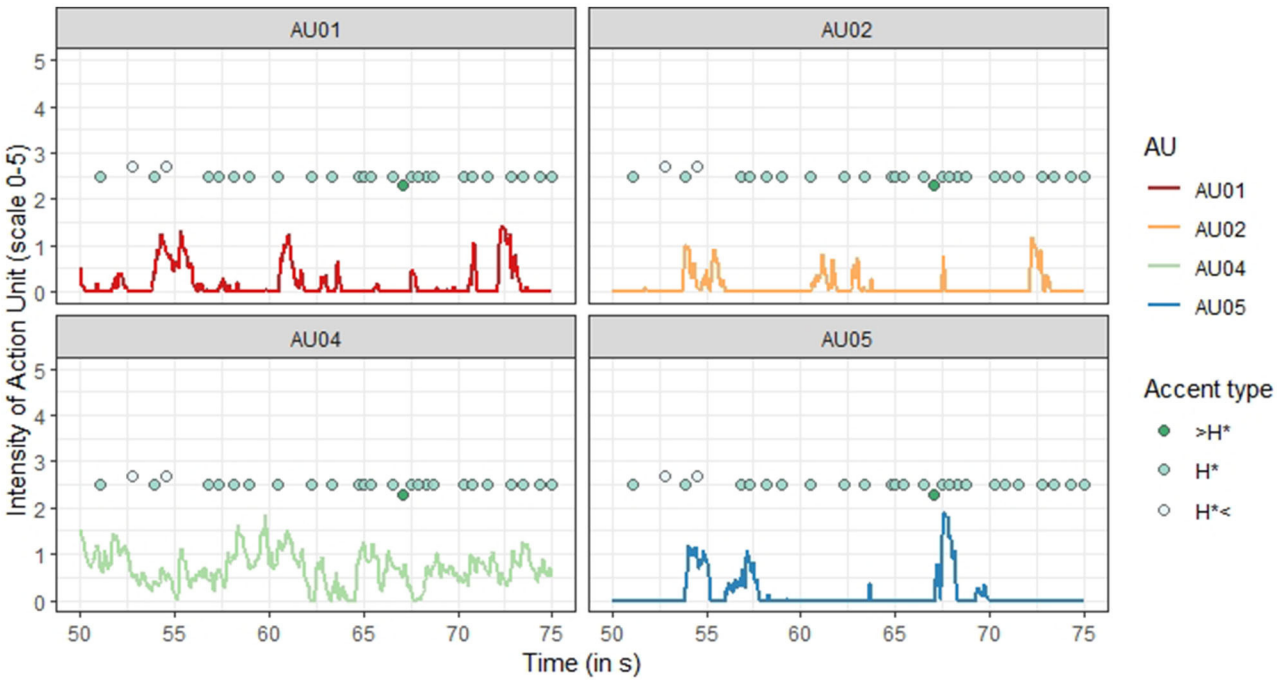


FIGURE 7 | Example of the occurrence of Action Units and pitch accents with their position relative to the prominent syllable. The intensity of each of the four Action Units (scale of 0–5) is shown for the first seconds 50 to 75 of the speech material of speaker LP. The pitch accents and their type are super-imposed, the y-axis does not refer to the accent type directly, but late peaks in the syllable after the prominent syllable (H*<) are placed higher than pitch peaks within the prominent syllable (H*), which in turn are placed higher than early peaks in the syllable before (>H*).

TABLE 5 | Number of usable contour pairs for each Action Unit.

Action unit	Contours
AU01	284
AU02	191
AU04	219
AU05	111

cases where the missing values fell between existing values, the missing values were linearly interpolated. If the missing values were at the beginning or the end of the sample and could not be interpolated, those time points were deleted from the sample, meaning that some samples were shorter than 1 s. In order to ensure that the samples were still long enough to be meaningful, only samples with at least 19 remaining time points were kept in the analysis. This resulted in 393 pitch accents for which a usable pitch contour existed.

Since we were particularly interested in the coordination of eyebrow movements, we also excluded cases in which the Action Unit value had a non-zero value in fewer than 10 of the datapoints. Since there were 4 tested AUs, this meant that a particular pitch accent unit might not appear in the analysis for, say, AU01, while still appearing in the analysis for AU04, if there were non-zero values for AU04 but not for AU01 for that pitch accent. The total number of contour pairs available for each Action Unit is shown in **Table 5**.

The FPCA analysis was carried out separately for each Action Unit, so each analysis involved a pair of contours: i.e., the F0 contour and the contour for one AU.

The analysis requires that all contours have the same duration. Since this is not the case due to the removal of some time points, a time-normalization must be carried out: the mean duration of all of the contours is calculated, and all contours are then normalized to this mean duration by either spreading out pitch points (to lengthen a contour) or moving them closer together (to shorten a contour). Thus the contours used as input for the analysis all have the normalized duration of 1 s, even if points were removed originally. It is possible to adjust the time-normalization process using additional time landmarks, e.g., segment or syllable onsets or offsets, but since the segmental content of these data were essentially random, no landmark registration was carried out for the current analysis.

Following this time normalization, a smoothing process is carried out. All of the contours are re-described using a common B-spline basis. A B-spline is a "piecewise" function, which is continuous at specified locations called knots. By linearly specifying the B-spline chunks, the intonation contours can be re-represented as smoothed functions. The number of knots as well as an additional smoothing parameter λ both impact the degree to which the contours are smoothed. There are several methods of determining optimal values for the number of knots and λ; the current study used a Generalized Cross-Validation method, which creates a visual representation of the relative amount of

error at different numbers of knots and λ values. Lower values for number of knots, and higher values for λ, lead to smoother curves; the degree of smoothness that is optimal for any dataset is in this case a result of the analyst's decision, carried out on the basis of visual inspection of the data. Once the parameters for smoothing have been determined, all of the contours are smoothed using the same basis.

Functional PCA results in a compact description of the main shape variations (or Principal Components, PCs) that are present within a dataset of curves; in this case, the set of F0 contours plus the set of eyebrow movement contours. Once the PCs are identified, each original curve is associated with a PC score, which quantifies where a specific curve is located in the continuum of each shape variation described by a PC.

Once the PCs are calculated, one output of the script is a set of visualizations of contours with a range of positive and negative values for each PC. The value for the PC can be thought of as a coefficient to the equation representing the contour, starting from the mean of all contours. The output for each PC also includes an account of how much variation in the data the PC accounts for.

Since PC scores are numbers, they can be easily used in further analyses, including being correlated with manual labels, as has been shown by Zellers et al. (2010), Gubian et al. (2015). Since they are directly related to the shape of the contours, they provide an objective method of identifying consistency and variability across a set of contours.

In the current study, in addition to the visual output of the PC analysis, we tested the output PC values against the prominence of the pitch accent as well as the identity of the speaker (to look for individual differences in alignment of eyebrow movement and F0).

Previous literature reported temporal coordination between eyebrow movements and F0 movements, and specifically, that raising of the eyebrows precedes the production of a pitch accent in speech. We therefore hypothesized that we would find peaks in eyebrow movement in the vicinity of the F0 peaks, most likely preceding them, such that the eyebrow contour would peak and then be in a fall (or already complete) by the time of the F0 peak. We also hypothesized that this pattern might be stronger in higher-prominence pitch accents, and that different speakers might show different patterns of F0-eyebrow coordination.

4.2. Results

The pitch contours and eyebrow movement contours were smoothed as described above, using $\log(\lambda) = -8$, and 12 knots. Since the results for all four AUs were substantially similar, the results for AU01, the inner brow raisers, are reported throughout as representative results. The first three PCs are shown in **Figure 8**.

There are several features of interest to be observed in **Figure 8**. First is the shape of the mean contours, shown in black. For the F0 contours (top), a clear peak is visible; this is to be expected, given that the data were selected on the basis of the presence of an F0 peak. However, in the eyebrow data, the mean contour is relatively flat. Some slight dipping can be observed at the left and right edges, but due to the time-warping, these locations must be considered as somewhat less reliable than locations in the middle of the contour. The lack of a peak in the mean contour for the eyebrow movements indicates that, on average, there is no eyebrow peak associated with the F0 peak (at least not in the time window used in this analysis). Thus these data do not provide support for a hypothesis proposing an automatic synchronization of the eyebrow raising movement with an F0 peak. However, since the mean value for the eyebrow contours falls between 2.5 and 5, the eyebrows (specifically in the case of AU01, the inner brow raisers) are on average relatively far away from their rest position while F0 is raised. It is important to keep in mind that this dataset excluded cases where no eyebrow movement at all was detected (i.e., a flat eyebrow contour with the AU value at 0) but did include cases where the AU value was even minimally above zero, so the mean of the contours suggests that, when eyebrow movements are present with F0 peaks, they are relatively intensely expressed, at least in terms of the way the AU is calculated (cf. Section 3.3).

The next feature of interest is the coordination of eyebrow raising and F0 in the first two PCs. The first two PCs account for approximately 70% of the variability in the data (PC1: 51.2%; PC2: 20.3%). The first PC accounts for matched variability: when F0 is higher, the eyebrows are also higher, aligning with our expectations. Thus, for a high value of PC1 (red contours), F0 and the AU value are both high). However, the second PC contributes an opposite effect: when F0 is higher, the eyebrows are lower. Thus, for a low value of PC2 (blue contours), F0 is high but the eyebrows are not raised. Although PC1 accounts for more variability in the data than PC2, PC2 could have an effect that essentially partially "cancels out" the effect of PC1, if the PC2 value for a given contour is sufficiently low while the PC1 value is not too high.

PC3 accounts for only 8.3% of the variability in the data, but shows a pattern more along the lines of our expectations; that is, a slope (either rising if PC3 is positive, or falling if PC3 is negative) in the eyebrow movement that is temporally close to the F0 peak, though not associated with a relatively higher or lower F0 peak. Although the variability accounted for is small, if there is a significant association of PC3 with another predictive variable, this could still be meaningful. However, as seen in **Figure 9**, this is not the case. Although in some cases, e.g., DH with prominence 2, LS or MF with prominence 1, or CB with prominence 3, there appears to be a difference in behavior, these visual differences rest on a lack of measurement points in these categories (only 2–3 datapoints) and thus do not attain statistical significance in a linear mixed model. No linear mixed model on any of the three PCs or combinations thereof found a significant relationship between the PCs and the prominence level. When the individual speakers were taken as predictors instead of as random factors in the model, individual speakers were occasionally found to differ significantly for one specific prominence value. Since this was not consistent across speakers or prominence levels, we conclude that, at least in the current data, there is no evidence for a specific systematic timing relationship between the F0 peak and any possible (though unidentified) peak in eyebrow movement activity.

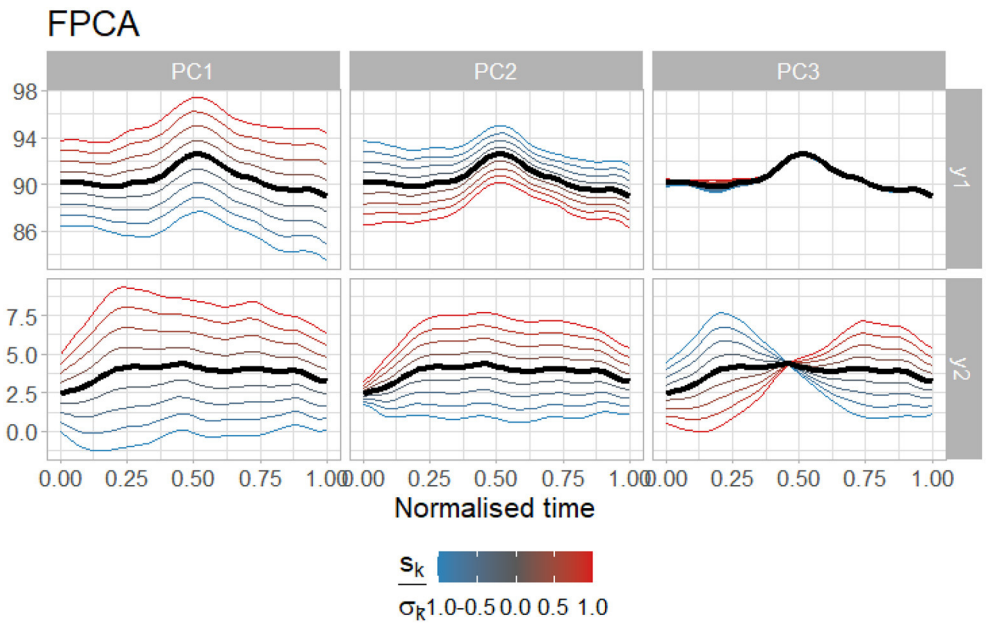


FIGURE 8 | The first three PCs resulting from the FPCA analysis of F0 with AU01. The top row (y1) shows the variability in F0, and the bottom row (y2) shows the variability in eyebrow movements. The black line shows the mean of all contours, while the red lines show contours with a high positive value for the given PC, and the blue lines show contours with a low negative value for the given PC. Note that the red and blue lines do not show density of the distributions of the PCs, and thus in some cases give non-real values (e.g., values below 0 or above 5 for y2, the eyebrow measurements.).

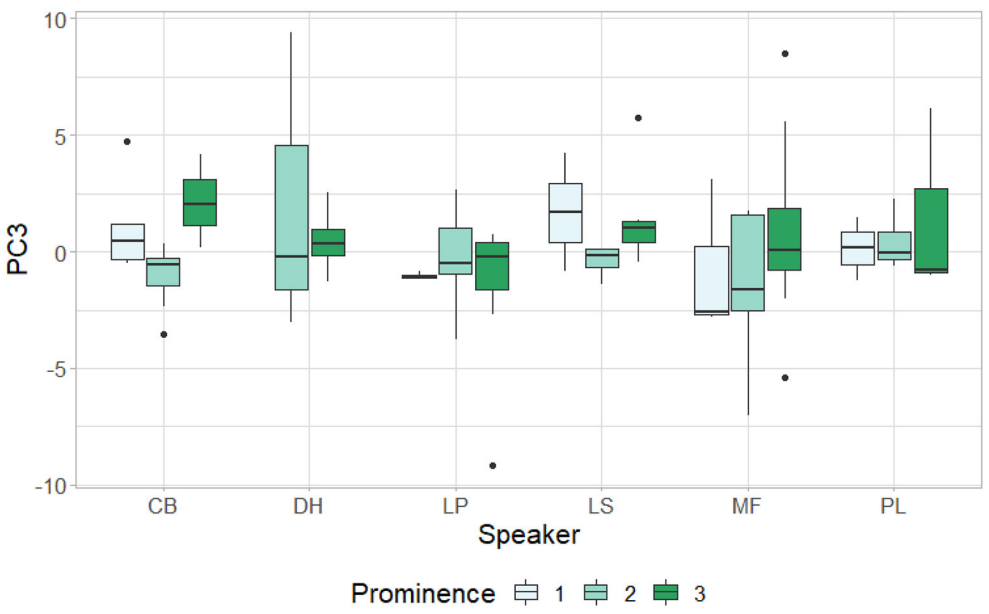


FIGURE 9 | Values of the third PC for each speaker across prominence levels.

5. DISCUSSION

Our first research question was if there is a correlation between the movement of the eyebrow (in terms of height for the eyebrow landmarks, the distance between eyebrow and eye landmarks, and intensity for AUs) and pitch height at any given moment in the analyses. We hypothesized that there will be correlations between pitch height and eyebrow movement suggesting that the eyebrows tend to be higher when the pitch was higher as well. This is exactly

what we found, at least for some of the landmarks and Action Units.

All speakers had at least one significant correlation with an AU. The Action Unit that had a significant—positive—correlation for all speakers was AU05, the upper lid raiser. This correlation suggests that the higher the pitch is, the stronger the raising of the upper lid and therefore the eyebrow. A widening of the eyes goes along with AU05 and should be investigated in future studies.

For the eyebrow landmarks we found that each speaker had at least three significant correlations which were all negative. The only two eyebrow landmarks that have significant correlations for all speakers are the middle point on the right eyebrow and the point to the right of the middle point. However, keep in mind that the lower the pixel value of the y-coordinate that is measured by OpenFace is, the higher the landmark appears in the video. Therefore, significant *negative* correlations also suggest that the higher the pitch is, the higher is the eyebrow in the video frame. Overall, the three speakers from North America in the current sample have stronger correlations between pitch height and eyebrow height than the speakers from England. This is a result that should be investigated further with a larger speaker sample. However, it might suggest that speakers from North America speak more expressively (= higher and therefore more prominent pitch peaks) and use their eyebrows more intensively (= higher eyebrow raises). This could fit some stereotypes between American and British cultures, which "attribute a more extrovert, expressive style and energetic manifestation of feelings to Americans than to the British, who are widely considered to be relatively more reserved" (Lewandowska-Tomaszczyk and Wilson, 2021, p. 262).

In order to corroborate that the landmark measurements actually represent eyebrow movements, we also used distance measures between eyebrow and eye landmarks as a more direct measure for eyebrow movement, as the distance between upper eyelid and eyebrow should not be massively affected by head movements. Two speakers (DH and LP) had only few or no significant correlations between pitch and distance measures. Only the correlations between the two outer landmark distances and pitch were significant for speaker DH, who also had only three significant (weak) correlations between the eyebrow landmarks and pitch. This may suggest that this particular speaker did not use the eyebrows much, especially in combination with pitch movements, and if he does it seems to be mostly with the outer ends of the eyebrows. For speaker LP, all correlations between eyebrow landmarks and pitch were significant, but none of the distance measures were. That was an unexpected result, as it was assumed that the distance measures would in a way confirm the position of the landmarks. It seems reasonable to assume that for this particular speaker, the correlations between the position of the eyebrow landmarks and pitch are more severely affected by head position in the video frame or head movements, so that it is not necessarily the *eyebrows* that were raised. For future studies we therefore suggest—inspired by Yunus et al. (2021)—to use both the landmarks and calculate distances to corroborate the results.

This is now also a concern for the other four speakers in the sample. However, the other four speakers had significant correlations for at least four out of the six distance measures with pitch as well as at least nine significant out of ten correlations between eyebrow landmarks and pitch. That means that the movements captured by the landmarks seem to actually be movements of the eyebrows, at least with a higher certainty. All four speakers (CB, LS, MF, and PL) have significant positive correlations between pitch and the distance from the center of each eyebrow to the respective eye landmark. That suggests that—at least for these four speakers—the center of the eyebrows is raised when the pitch is higher. That would also fit the significant positive correlations with AU05, the upper lid raiser: when the upper lid is raised, especially the center point of the eyebrows moves away from the eyes. But there are, again, strong speaker-specific differences between the eyebrow movements and pitch height. The two speakers who did not have significant correlations between the distances between eyebrows and eyes and pitch, or just at very few points were two of the three speakers from England, while all speakers from North America had significant correlations between several distances and pitch. This could also point toward more expressivity for American speakers and perhaps a tendency for more reservedness for speakers from England (cf. Lewandowska-Tomaszczyk and Wilson, 2021).

In general, there seems to be no specific difference in eyebrow and pitch use depending on the acoustic charisma score that was calculated for the speakers (cf. Niebuhr et al., 2017; Niebuhr, 2021). Rather, all speakers had fairly similar scores, but strong differences between correlations of eyebrow movements and pitch height. That suggests that acoustic charisma and eyebrow movements are not necessarily tied together, though it is also not dis-proven given the small dataset and the study not being directly aimed at investigating correlations with the charisma scores. However, one possible reading could be that there might be a trade-off between verbal and non-verbal charisma, as speaker LS has—with 45—the lowest charisma score but at the same time the strongest correlations between eyebrow movements and pitch. It might therefore be that she makes up for her lower verbal charisma with facial charisma. Similar results have been found for hand gestures (see, for example, Hiroyuki and Rathcke, 2016).

We also carried out binomial regression models to find out whether the presence of a pitch accent of a specific prominence level (weak, strong, emphatic) or accent type (early, medial, late) can predict the presence or absence of an Action Unit connected to the eyebrows. We found that the presence of a pitch accent with a particular prominence level does not predict the presence or absence of an AU. However, we did find that the presence of a medial pitch accent (H*, occurring *within* the prominent syllable) can in fact predict the presence of an inner brow raiser (AU01). This suggest some kind of co-occurrence between medial pitch accents and AU01, and while medial pitch accents are by far the most frequent accent type (as they are the default), the presence of AU02 (outer brow raiser) and AU04 (brow lowerer) could not be predicted by the presence of a medial pitch accent. So even though medial pitch accents are the default, they seem to be more connected to movement of the inner brows than

of the outer brows. The presence of a medial pitch accent also predicts the presence of an upper lid raiser movement (AU05). The presence of AU05 is also predicted by the presence of a late pitch accent (H*<, occurring in the syllable *following* the prominent syllable). That may suggest that the upper lid raiser could be more prominent in general, as it is more likely to co-occur with non-early and therefore more prominent pitch accents because of their interpretation as "unusual occurrence" (Gussenhoven, 2002, p. 6). It is also the only Action Unit that was significantly positively correlated with pitch height for all six speakers in the sample, further suggesting the connection to prominence. AU05 raises the entire lid and therefore moves the eyebrows up entirely, which can convey a look of surprise, also fitting the "unusual occurrence" interpretation. This tends to go together with a widening of the eyes. Eyebrow raising and widening are also used to signal overarching prosody in sign languages (see Nespor and Sandler, 1999 for Israeli Sign Language), which may also hint at an interplay between the two gestures and warrant further analysis in terms of prominence.

The FDA analysis, investigating co-occurring movement between the eyebrows and F0 peaks, did not provide a conclusive result, but opened up several possibilities for future research. We hypothesized that individual eyebrow movements might be closely synchronized in time with F0 peaks, and that this synchronization might be stronger at higher prominence levels, or vary among individual speakers. However, we found no evidence supporting either of these hypotheses. This contrasts with previous reports from Flecha-García (2010), who suggests that there is a very close temporal alignment between at least the raising movement of the eyebrows and the location of an F0 peak. It is important to keep in mind that the methodology used was different, and that the previous studies used manual annotations rather than automatically extracted AUs. Since the current study could only analyze one AU at a time in relation to a pitch accent (increasing the number of dimensions of the FDA would have substantially increased processing time as well as reducing the dataset to those items where more than one AU was active), it is possible that the manual annotations were able to capture complex movements that were not reliably relayed by the OpenFace data. Conversely, manual annotators might have been unconsciously influenced by body movements or other aspects of the video signal which the automatic method employed here avoided. It might also be the case, as suggested by a reviewer, that the automatic method was sensitive to movements that were under a threshold of visual perceptibility to human raters. Since our analysis included all non-zero values for the AUs, it is possible that eyebrow movements were included that would not have been classified as eyebrow movements by a human rater. Thus, while one might claim that differences in the results are due to differences in validity of the data extraction method, it is difficult to evaluate which method might have produced more reliable data.

The current data also differ in genre from Flecha-García (2010), who uses dialogues in which participants sit face to face, a situation which was likely to encourage non-verbal signaling to the interlocutor. In the YouTube data, although the speakers are interacting with an audience, this audience is normally not visible to them, and this may have a damping effect on signaling. Previous research also found this: when interlocutors were visible to each other, speakers "were more likely to use *non-redundant gestures*" than speakers who could not see their interlocutor (Bavelas and Healing, 2013, p. 77, see also Alibali et al., 2001). Gestures have likewise been found to be larger when interlocutors could see each other (Mol et al., 2011). This is likely also the case on YouTube where the audience is not visible, and could explain the lack of alignment between eyebrow movements and pitch accents. The effect on other types of gestures (manual gestures, head nods, etc.) would be interesting to investigate in this light in the future. Conversely, perhaps "YouTube Face", a conventionalized method or set of methods of visually signaling engagement or emphasis, will turn out to be a visual corollary to the "YouTube Voice" which has developed in this particular speaking context.

Another crucial feature of the current study that differs from previous analyses is that it used the pitch accent as a reference point for identifying the time window in which we investigated eyebrow movement. If only the eyebrow raise, and not the subsequent fall, is closely timed to the pitch accent, then it is possible that the window of 500 ms excluded the time in which the eyebrow raise occurred; Flecha-García (2010, p. 549) reports eyebrow movements starting within ±1 s of the pitch accent. However, given that pitch accents in the current data tended to be between 600 and 800 ms apart, this seems unlikely, if there was a one-to-one correlation between eyebrow raises and F0 peaks. A one-to-one correlation is in any case not to be expected due to the extreme degree of eyebrow movement that would be involved (nor is such a one-to-one relationship found in previous work), although given the association with increased prominence found in previous studies, it might have been reasonable to find this association with the prominence level 3 accents, even if not for levels 1 and 2.

The findings in the current data appear to be more in line with a model of eyebrow movement where the eyebrows are raised and then perhaps held in the raised position over the course of several pitch accents. This would be consistent with the mean value of the eyebrow contours falling clearly in the upper range of activity for the AU, while not having a clear contour form, as this might be lost in the calculation of the average contour. Since this study investigated all eyebrow movements identified by OpenFace, it is possible that a confound arose between rapid and slow eyebrow movements, i.e., those having rhythmic or prominence functions vs. those having semiotic functions (Guaïtella et al., 2009). A longer holding of the eyebrow movements might correlate better with a "structuring" function for eyebrow movements, which Flecha-García (2010) proposes as being of similar importance to emphasis. A future investigation using FDA could investigate the shapes of eyebrow movement contours: do local eyebrow peaks analogous to F0 peaks tend to arise? Or are long plateaus of raised eyebrows more likely? Possibly excluding the F0, multidimensional analyses could also look at multiple AUs simultaneously. As suggested by the work of Kim et al. (2014), a measure of peak velocity could be a better indicator for identifying relevant eyebrow movements.

In general, as we have hypothesized, we found both speaker-specific differences and tendencies for origin differences. That suggests that the use of eyebrow movements and pitch accents is not universal, but depending on the preferences of the specific speaker. Additionally, it seems that eyebrow movements and pitch accents are more closely tied for North American speakers than for speakers from England, perhaps hinting at a difference in expressivity between the two cultures (Lewandowska-Tomaszczyk and Wilson, 2021). Gender differences were not found in this particular sample. While this is a generalization and only first impressions, we see these tendencies for our sample that warrant further investigations with a larger sample. Nevertheless, this exploration revealed that there may be origin-based differences in how eyebrow movements and pitch accents play together.

Overall, we found correlations between the height of eyebrow movements and the height of the pitch contour at the same measuring points. That suggests, as other studies have shown (e.g., Flecha-García, 2010; Ambrazaitis and House, 2017), that eyebrow movements and intonation are somewhat connected. However, unlike other studies (e.g., Swerts and Krahmer, 2010), we did not find evidence for a difference in eyebrow movement depending on the prominence level of a pitch accent. It was also not likely to have pitch accents at the same time as Action Units, at least for some of the Action Units. That suggests that pitch accents may be connected with certain *types* of eyebrow movements (primarily in this sample: inner brow raiser and upper lid raiser), and only for specific pitch accent positions relative to the prominent syllable (medial or late). There are, however, hints to a cultural difference between speakers from North America and England that needs further analysis. The YouTubers in the sample frequently use prominence strategies like emphatic and late pitch accents, and the visual component is integral in a video. Our results of a connection of pitch accents and eyebrow movements are tentatively positive, but require further investigations.

Future studies should expand the sample of speakers, both from North America and from England. This study was also only the first step and should be enhanced further by running perception experiments with visual, audio, and audio-visual cues to determine to what degree the eyebrows actually have an effect on the perception of expressiveness, charisma, and the overall personality of the speaker. Additionally, future investigations should also incorporate head nods and other head movement gestures. That might add another relevant level to the question of gesture-pitch alignment and could lead to a more complete picture. The information status of the verbal items could be investigated, and if there are certain co-occurrences with pitch peaks and eyebrow movements. Finally, future studies could also investigate if there is a difference in amount of AUs in the beginning, middle and end of a talk or video: especially on YouTube it is important to grab an audience from the very beginning. That might suggest that AUs are more frequent in the beginning of a video to get as much expressiveness as possible into the first section of the video to engage the viewer.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

Ethical review and approval was not required for the current study in accordance with the local legislation and institutional requirements. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

AUTHOR CONTRIBUTIONS

SB contributed to hypothesis development, annotated the pitch accents and prominence levels, carried out the OpenFace analyses and the data processing and statistics for Section 3. MZ contributed to hypothesis development, carried out the data processing and statistics for Section 4, and contributed to writing the paper. All authors contributed to the article and approved the submitted version.

FUNDING

Part of the work of SB was supported financially by the Federal State Grant Schleswig-Holstein (awarded between 2018 and 2021). We acknowledge financial support by DFG within the funding programme Open Access-Publikationskosten.

ACKNOWLEDGMENTS

The authors would like to thank Michele Gubian for the support with the Functional Data Analysis, and Oliver Niebuhr for the calculation of the PASCAL scores. Thank you also to the reviewers as well as Kathrin Feindt for their helpful comments and suggestions. Part of the work carried out was inspired by discussions at meetings of the network Gestures and Head Movements in Language (Independent Research Fund Denmark, 9055-00004B, PI: Prof. Patrizia Paggio).

REFERENCES

Akker, E., and Cutler, A. (2003). Prosodic cues to semantic structure in native and nonnative listening. *Bilingual. Lang. Cogn.* 6, 81–96. doi: 10.1017/S1366728903001056

Alibali, M. W., Heath, D. C., and Myers, H. J. (2001). Effects of visibility between speaker and listener on gesture production: some gestures are meant to be seen. *J. Mem. Lang.* 44, 169–188. doi: 10.1006/jmla.2000.2752

Ambrazaitis, G., and House, D. (2017). Multimodal prominences: exploring the patterning and usage of focal pitch accents, head beats and eyebrow beats in Swedish television news readings. *Speech Commun.* 95, 100–113. doi: 10.1016/j.specom.2017.08.008

Asano, Y., and Gubian, M. (2018). "Excuse meeee!!": (mis)-coordination of lexical and paralinguistic prosody in L2 hyperarticulation. *Speech Commun.* 99, 183–200. doi: 10.1016/j.specom.2017.12.011

Ballinger, C. (2017). *My Experience With Netflix*. Available online at: https://www.youtube.com/watch?v=KWM4DVRrdaA

Baltrusaitis, T., Zadeh, A., Lim, Y. C., and Morency, L.-P. (2018). "OpenFace 2.0: facial behavior analysis toolkit," in *2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018)* (IEEE), 59–66. doi: 10.1109/FG.2018.00019

Bates, D., Sarkar, D., Bates, M. D., and Matrix, L. (2007). *The lme4 Package*. R package version 2.

Bavelas, J., and Healing, S. (2013). Reconciling the effects of mutual visibility on gesturing: a review. *Gesture* 13, 63–92. doi: 10.1075/gest.13.1.03bav

Beck, J. (2015). *The Linguistics of 'YouTube Voice'*. Boston, MA: The Atlantic. Available online at: https://www.theatlantic.com/technology/archive/2015/12/the-linguistics-of-youtube-voice/418962/

Berger, S., Niebuhr, O., and Brem, A. (2020). "Of voices and votes: phonetic charisma and the myth of Nixon's radio victory in his first 1960 TV debate with Kennedy," in *An den Rändern der Sprache*, eds M. Elmentaler and O. Niebuhr (Frankfurt: Peter Lang), 109–145.

Berger, S., and Zellers, M. (2021). "Pitch accent position, peak height, and prominence level relative to accented vowel onset on YouTube," in *Proceedings of the 1st International Conference on Tone and Intonation* (Sønderborg), 137–141.

Beskow, J., Granström, B., and House, D. (2006). "Visual correlates to prominence in several expressive modes," in *Ninth International Conference on Spoken Language Processing* (Pittsburgh, PA: Citeseer), 1272–1275. doi: 10.21437/Interspeech.2006-375

Boersma, P., and Weenink, D. (2018). *Praat: Doing Phonetics by Computer, Version 6.0.37*. Available online at: https://www.fon.hum.uva.nl/praat/

Bolinger, D. (1972). Accent is predictable (if you're a mind-reader). *Language* 48, 633–644. doi: 10.2307/412039

Cannata, D., Redfern, S., and O'Hora, D. (2020). "OpenFaceR: developing an R package for the convenient analysis of OpenFace facial information," in *PSYCHOBIT* (Naples), 1–9.

Cavé, C., Guaitella, I., Bertrand, R., Santi, S., Harlay, F., and Espesser, R. (1996). "About the relationship between eyebrow movements and Fo variations," in *Proceeding of Fourth International Conference on Spoken Language Processing. ICSLP'96* (Philadelphia, PA: IEEE), 2175–2178. doi: 10.1109/ICSLP.1996.607235

Cheng, C., and Gubian, M. (2011). "Predicting Taiwan Mandarin tone shapes from their duration," in *Proceedings of Interspeech 2011* (Florence), 1073–1076. doi: 10.21437/Interspeech.2011-403

Cheng, C., Xu, Y., and Gubian, M. (2010). "Exploring the mechanism of tonal contraction in Taiwan Mandarin," in *Proceedings of Interspeech 2010* (Makuhari), 2010–2013. doi: 10.21437/Interspeech.2010-569

Ekman, P., and Friesen, W. V. (1978). *Facial Action Coding System: A Technique for the Measurement of Facial Movement*. Palo Alto, CA: Consulting Psychologists Press.

Esteve-Gibert, N., and Prieto, P. (2013). Prosodic structure shapes the temporal realization of intonation and manual gesture movements. *J. Speech Lang. Hear. Res.* 56, 850–864. doi: 10.1044/1092-4388(2012/12-0049)

Fischbach, M. (2018). *Let's Be Completely Honest*. Available online at: https://www.youtube.com/watch?v=T6cdM1kubk4

Flecha-García, M. L. (2010). Eyebrow raises in dialogue and their relation to discourse structure, utterance function and pitch accents in English. *Speech Commun.* 52, 542–554. doi: 10.1016/j.specom.2009.12.003

FreeTime (2021). *Formatfactory*.

Granström, B., House, D., and Lundeberg, M. (1999). "Prosodic cues in multimodal speech perception," in *Proceedings of the International Congress of Phonetic Sciences* (San Fransisco, CA: ICPhS99), 655–658.

Green, A. (2015). *What is 'YouTube Voice?' A Linguist Breaks It Down*. Mental Floss.

Guaïtella, I., Santi, S., Lagrue, B., and Cavé, C. (2009). Are eyebrow movements linked to voice variations and turn-taking in dialogue? An experimental investigation. *Lang. Speech* 52, 207–222. doi: 10.1177/0023830909103167

Gubian, M., Torreira, F., and Boves, L. (2015). Using functional data analysis for investigating multidimensional dynamic phonetic contrasts. *J. Phonet.* 49, 16–40. doi: 10.1016/j.wocn.2014.10.001

Gubian, M., Torreira, F. J., Strik, H., and Boves, L. W. (2009). "Functional data analysis as a tool for analyzing speech dynamics: a case study on the French word C'était," in *Proceedings of Interspeech 2009* (Brighton), 2199–2202. doi: 10.21437/Interspeech.2009-626

Gussenhoven, C. (2002). "Intonation and interpretation: phonetics and phonology," in *Speech Prosody 2002, International Conference* (Aix-en-Provence), 1–10.

Hagi, S. (2017). *The Rise of the 'YouTube Voice' and Why Vloggers Want It to Stop*. Available online at: https://www.vice.com/en/article/aepn94/the-rise-of-youtube-voice-and-why-vloggers-want-it-to-stop

Hiroyuki, T., and Rathcke, T. (2016). "Then, what is charisma?: the role of audio-visual prosody in L1 and L2 political speeches," in *Phonetik und Phonologie im deutschsprachigen Raum* (Munich), 1–3.

Howell, D. (2017). *Daniel and Depression*. Available online at: https://www.youtube.com/watch?v=Wp2TUPo5W0c

Ip, M. H. K., and Cutler, A. (2021). "In search of salience: focus detection in the speech of different talkers," in *Language and Speech* (New York, NY: Sage Publications), 1–31. doi: 10.1177/00238309211046029

Jennings, R. (2021). *How Should An Influencer Sound*? Available online at: https://www.vox.com/the-goods/2021/7/13/22570476/youtube-voice-tiktok-influencer-sound

Kim, J., Cvejic, E., and Davis, C. (2014). Tracking eyebrows and head gestures associated with spoken prosody. *Speech Commun.* 57, 317–330. doi: 10.1016/j.specom.2013.06.003

Kohler, K. J., and Niebuhr, O. (2007). "The phonetics of emphasis," in *Proc. 16th ICPhS, Saarbrücken* (Saarbrücken), 2145–2148.

Krivokapic, J., Tiede, M., and Tyrone, M. E. (2015). "A kinematic analysis of prosodic structure in speech and manual gestures," in *ICPhS* (Glasgow).

Kügler, F., and Baumann, S. (2019). *Annotationsrichtlinien DIMA; version 4.0*. DIMA Website. Available online at: http://dima.uni-koeln.de/wp-content/uploads/2019/02/DIMA-Annotationsrichtlinien.V4.0.pdf

Kügler, F., Baumann, S., Andreeva, B., Braun, B., Grice, M., Neitsch, J., et al. (2019). "Annotation of German intonation: DIMA compared with other annotation systems," in *Proceedings of the International Congress of Phonetic Sciences* (Melbourne, VIC: ICPhS), 1–4.

Kügler, F., Smolibocki, B., Arnold, D., Baumann, S., Braun, B., Grice, M., et al. (2015). "DIMA: annotation guidelines for German intonation," in *ICPhS 2015: 18th International Congress of Phonetic Sciences* (Glasgow).

Ladd, D. R. (2008). *Intonational Phonology*. Cambridge, UK: Cambridge University Press. doi: 10.1017/CBO9780511808814

Lester, P. (2018). *Why I Went to Hospital*. Available online at: https://www.youtube.com/watch?v=IDrBKN9Lgk4

Lewandowska-Tomaszczyk, B., and Wilson, P. A. (2021). "Expressive and reserved cultural linguistic schemas: British and American pride clusters," in *Cultural Linguistics and World Englishes*, eds M. Sadeghpour and F. Sharifian (Singapore: Springer), 261–293. doi: 10.1007/978-981-15-4696-9_13

Loehr, D. P. (2004). *Gesture and intonation* (Ph.D. thesis). Georgetown University, Washington, DC, United States.

Loehr, D. P. (2012). Temporal, structural, and pragmatic synchrony between intonation and gesture. *Lab. Phonol.* 3, 71–89. doi: 10.1515/lp-2012-0006

Lohfink, G., Katsika, A., and Arvaniti, A. (2019). "Variability and category overlap in the realization of intonation," in *Proceedings of ICPhS 2019* (Canberra, ACT).

McNeill, D. (2008). *Gesture and Thought*. Chicago, IL: University of Chicago Press.

Mo, Y. (2008). "Duration and intensity as perceptual cues for naive listeners' prominence and boundary perception," in *Proceedings of the 4th Speech Prosody Conference* (Campinas), 739–742.

Mol, L., Krahmer, E., Maes, A., and Swerts, M. (2011). Seeing and being seen: the effects on gesture production. *J. Comput. Mediat. Commun.* 17, 77–100. doi: 10.1111/j.1083-6101.2011.01558.x

Namba, S., Sato, W., Osumi, M., and Shimokawa, K. (2021a). Assessing automated facial action unit detection systems for analyzing cross-domain facial expression databases. *Sensors* 21, 4222. doi: 10.3390/s21124222

Namba, S., Sato, W., and Yoshikawa, S. (2021b). Viewpoint robustness of automated facial action unit detection systems. *Appl. Sci.* 11, 11171. doi: 10.3390/app112311171

Nasir, M., Jati, A., Shivakumar, P. G., Nallan Chakravarthula, S., and Georgiou, P. (2016). "Multimodal and multiresolution depression detection from speech and facial landmark features," in *Proceedings of the 6th International Workshop on Audio/Visual Emotion Challenge* (Amsterdam), 43–50. doi: 10.1145/2988257.2988261

Nespor, M., and Sandler, W. (1999). Prosody in israeli sign language. *Lang. Speech* 42, 143–176. doi: 10.1177/00238309990420020201

Niebuhr, O. (2010). On the phonetics of intensifying emphasis in German. *Phonetica* 67, 170–198. doi: 10.1159/000321054

Niebuhr, O. (2021). "Advancing higher-education practice by analyzing and training students' vocal charisma: evidence from a Danish field study," in *7th International Conference on Higher Education Advances* (Valencia), 743–751. doi: 10.4995/HEAd21.2021.12827

Niebuhr, O., Brem, A., Michalsky, J., and Neitsch, J. (2020). What makes business speakers sound charismatic? A contrastive acoustic-melodic analysis of Steve Jobs and Mark Zuckerberg. *Cadernos de Linguística* 1, 1–40. doi: 10.25189/2675-4916.2020.v1.n1.id272

Niebuhr, O., Brem, A., and Tegtmeier, S. (2017). Advancing research and practice in entrepreneurship through speech analysis-from descriptive rhetorical terms to phonetically informed acoustic charisma profiles. *J. Speech Sci*, 6, 3–26. doi: 10.20396/joss.v6i1.14983

Niebuhr, O., Voße, J., and Brem, A. (2016). What makes a charismatic speaker? A computer-based acoustic-prosodic analysis of Steve Jobs tone of voice. *Comput. Hum. Behav*. 64, 366–382. doi: 10.1016/j.chb.2016.06.059

Pentland, L. (2017). *I'm So Sorry | IWD2017*. Available online at: https://www.youtube.com/watch?v=sezwogn4zOE

Ramsay, J., and Silverman, B. (2005). *Functional Data Analysis*. New York, NY: Springer. doi: 10.1007/b98888

RStudio Team (2021). *RStudio: Integrated Development Environment for R*. Boston, MA: RStudio, PBC.

Singh, L. (2017). *We Need to Have an Honest Talk*. Available online at: https://www.youtube.com/watch?v=KjK81YmQEuY

Sridhar, V. K. R., Nenkova, A., Narayanan, S., and Jurafsky, D. (2008). "Detecting prominence in conversational speech: pitch accent, givenness and focus," in *Proceedings of Speech Prosody* (Campinas: Citeseer), 456.

Swerts, M., and Krahmer, E. (2008). Facial expression and prosodic prominence: effects of modality and facial area. *J. Phonet*. 36, 219–238. doi: 10.1016/j.wocn.2007.05.001

Swerts, M., and Krahmer, E. (2010). Visual prosody of newsreaders: effects of information structure, emotional content and intended audience on facial expressions. *J. Phonet*. 38, 197–206. doi: 10.1016/j.wocn.2009.10.002

Turco, G., Gubian, M., and Schertz, J. (2011). "A quantitative investigation of the prosody of Verum Focus in Italian," in *Proceedings of Interspeech 2011* (Florence), 961–964. doi: 10.21437/Interspeech.2011-393

Vainio, M., and Järvikivi, J. (2006). Tonal features, intensity, and word order in the perception of prominence. *J. Phonet*. 34, 319–342. doi: 10.1016/j.wocn.2005.06.004

Wagner, P., Malisz, Z., and Kopp, S. (2014). Gesture and speech in interaction: an overview. *Speech Commun*. 57, 209–232. doi: 10.1016/j.specom.2013.09.008

Wichmann, A., House, J., and Rietveld, T. (1997). "Peak displacement and topic structure," in *Intonation: Theory, Models and Applications* (Athens), 329–332.

Yunus, F., Clavel, C., and Pelachaud, C. (2021). "Sequence-to-sequence predictive model: from prosody to communicative gestures," in *Digital Human Modeling and Applications in Health, Safety, Ergonomics and Risk Management. Human Body, Motion and Behavior. HCII 2021. Lecture Notes in Computer Science, Vol 12777*, ed V. G. Duffy (Cham: Springer), 355–374. doi: 10.1007/978-3-030-77817-0_25

Zellers, M., Gubian, M., and Post, B. (2010). "Redescribing intonational categories with Functional Data Analysis," in *Proceedings of Interspeech 2010* (Makuhari), 1141–1144. doi: 10.21437/Interspeech.2010-361

Zellers, M., Post, B., and D'Imperio, M. (2009). "Modeling the intonation of topic structure: two approaches," in *Proceedings of Interspeech 2009* (Brighton). doi: 10.21437/Interspeech.2009-315

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Berger and Zellers. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers | Frontiers in Psychology

BRIEF RESEARCH REPORT
published: 21 June 2022
doi: 10.3389/fpsyg.2022.911556

President Obama's Humble Face: An Authentic or a Socially Desirable Posturing? A Study on Reactions to Obama's Autobiographical Self-Disclosures

Alessia Mastropietro[1], Peter Bull[2], Francesca D'Errico[3], Isora Sessa[4], Stefano Migliorisi[4] and Giovanna Leone[4]*

[1] Department of Psychology, Sapienza University of Rome, Rome, Italy, [2] Department of Psychology, University of York, York, United Kingdom, [3] Department of Education, Psychology, and Communication, University of Bari "Aldo Moro", Bari, Italy, [4] Department of Communication and Social Research, Sapienza University of Rome, Rome, Italy

Referring to the mainstream studies based on the personalization's hypothesis, which positively evaluates signals of dominance shown by leaders, the analysis of Obama's rhetoric stays a relevant exception. His risky recall, during his political talks, of his social difficulties as a child of a mixed couple was in fact one of the more surprising aspects of his success. Nevertheless, reactions to his autobiographical sharing were scarcely explored. Based on the idea that these self-disclosures signal his responsivity toward the audience of low social condition and can, therefore, be defined as a sign of humility, this research aims to test if coherence between Obama's words and his facial expressions of contempt, due to the seriousness of social injustices endured in his childhood, may influence the receivers' perception of such unexpected communication. Before reading a brief autobiographical sharing taken from a "Back-to-school" speech, a highly ritualized monolog the US President addresses each year to students, 175 Italian participants were presented with a photo of Obama displaying either an expression of contempt (taken from the video of the speech) or a neutral expression. Comparisons between self-assessments of perceptions and reactions of participants assigned to the two experimental conditions show that a facial expression of contempt, coherent with words describing his school difficulties, has been crucial for perceiving this humble political discourse as authentic and not as a simple socially desirable posturing. More studies seem to be needed, however, to understand how humble speech could enhance the positive face of leaders or backfire against them.

Keywords: humility, Obama, FACS, political speech, autobiographical social sharing, emotions, contempt

OPEN ACCESS

Edited by:
Willibald Ruch,
University of Zurich, Switzerland

Reviewed by:
Maria Giuseppina Pacilli,
University of Perugia, Italy
Pierre Gander,
University of Gothenburg, Sweden

****Correspondence:***
Alessia Mastropietro
alessia.mastropietro@uniroma1.it

Specialty section:
This article was submitted to Personality and Social Psychology, a section of the journal Frontiers in Psychology

Received: *02 April 2022*
Accepted: *19 May 2022*
Published: *21 June 2022*

Citation:
Mastropietro A, Bull P, D'Errico F, Sessa I, Migliorisi S and Leone G (2022) President Obama's Humble Face: An Authentic or a Socially Desirable Posturing? A Study on Reactions to Obama's Autobiographical Self-Disclosures. Front. Psychol. 13:911556. doi: 10.3389/fpsyg.2022.911556

INTRODUCTION

This article is based on the idea that expressions of humility can positively contribute to politicians' facework, meant to show a positive face to their audience (Bull and Fetzer, 2010), but only when specific situational and interactive conditions are fulfilled. In this study, we started to test this idea by first analyzing a genre of political communication that Bull and Fetzer (2010) defined as

a "monolog." In this specific situation, we assumed that the interactive condition leading to a good contribution of humility's expressions to the politician's positive face was the receivers' perception that this humble communication was sincere. However, before discussing the hypotheses tested in the study and the choice of the monolog analyzed, a brief review of studies on humility in political communication is in order.

Humility in Political Speech

Humility can be defined as a multimodal public stance interactively performed through verbal and non-verbal signals, whereby the person places themselves in a horizontal relationship to the interlocutor (D'Errico and Poggi, 2019). Therefore, a multimodal analysis seems to be the best methodological choice to observe and understand this kind of communication. In particular, the humble stance may be defined according to several aspects, namely, the tendency to express positive other-oriented emotions (e.g., empathy and compassion), the ability to regulate self-oriented emotions in socially acceptable ways (e.g., pride or excitement about one's accomplishments; Davis et al., 2013), and the orientation of showing responsiveness to others' needs (Fetzer and Bull, 2012). Considering recent studies on political personalization, a humble stance can seem useless, if not counteractive and incoherent, since it may be perceived as "unauthentic" (Luebke, 2020), especially in male politicians (Liu, 2016; D'Errico, 2019). Nevertheless, it could become very effective in conflictual situations, in which a humble stance could help speakers to better elaborate their own point of view.

Interestingly, a recent content analysis of common sense on humility (Weidman et al., 2018) highlighted a two-sided representation, which authors, respectively, named self-abasing humility and appreciative humility. The self-abasing humility, or modesty, motivates a speaker to hide from others and is associated with feelings of submissiveness, worthlessness, and traits, such as introversion and low self-esteem. Appreciative humility is instead associated with compassion, grace, and understanding, as well as with high status, high self-esteem, and agreeableness. The appreciative dimension of humility is related to behaviors, such as giving space to others' opinions, admitting one's own mistakes and gaps, and giving others' merits (Liu, 2016; Weidman et al., 2018), as it seems to be based on a deep self-awareness of both one's own strengths and limits (Tangney, 2000).

Based on such an appreciative definition of humility, recent studies explored the important role of emotional expressiveness in the perception of the politician's authenticity.

On the one hand, overt expression of emotions, such as when showing anger or sadness, helps politicians to emphasize the importance of the topic at hand (Scardigno et al., 2021). However, when the specific case of humble communication is at stake, different pragmatic effects were observed depending on the specific emotion shown. Male political leaders proved to be especially effective when conveying a moral message and showing an angry facial expression; but they were perceived negatively, or their message was supposed to be hypocritical, when showing a sad expression (D'Errico, 2019). On the contrary, humble communication of female politicians seemed to elicit positive evaluations of their competence and benevolence when they exhibited a sad facial display during a persuasive message conveying a moral gist (D'Errico et al., 2022). This difference could be accounted for if considering the stereotypical expectancies linked to the social role assigned to men and women. While men are expected to be dominant, and therefore ready to express emotions more extroverted and linked to activity, such as rage, women are expected to express emotions linked to introversion and passivity, such as sadness (Eagly, 1987, 2005). Emotions shown by leaders could, therefore, be judged either as spontaneous or faked, depending on their coherence with previous expectations held by receivers. Similarly, referring to the social dominance of the group from which leaders originated, leaders of dominant groups are expected to be more agentic than the dominated ones.

Based on this line of thought, it is interesting to observe that one main rhetorical strategy used consistently by Barack Obama, both when speaking as an incumbent and when playing the official role of US President, was to overtly refer to him being the son of a mixed married couple, i.e., a social origin illegal in some North American states at the time of his birth. We speculated (Leone et al., 2015, 2018b) that, choosing such an authentic self-presentation (Gunn, 2015), he acted in a way that Arendt (1978) defined as being an "aware pariah," i.e., an offspring of a seriously dominated group that shows pride and gratitude for his social origin instead that conceals it. According to the definitions of appreciative humility, we consider his unveiling of autobiographical memories of difficulties he encountered during his childhood and adolescence as a good instance of a humble stance. Moreover, we sustain that, when speaking to an audience of similar dominated social origins, this social sharing of autobiographical memories could also be taken as a sign of responsive humility.

The Case Study

Within this theoretical framework, the case study selected refers to a "Back-to-school" speech, delivered by the former US President Barack Obama on 8 September 2009, in front of students of low socioeconomic status at the Wakefield High School of Arlington, Virginia. The "Back-to-school" Presidential address is a highly ritualized monolog, where each American leader in charge is expected to encourage students, at the beginning of every year's lessons, to make their best to become competent citizens, aware that the future of their community depends on them. In a previous study (Leone et al., 2018a), we analyzed in-depth two Back-to-school speeches of Barack Obama, addressing students either of high or of low social status. Although in both speeches the President shared with students an autobiographical recall of serious difficulties of his adolescent times, a multimodal analysis highlighted that he expressed an emotion coherent with his verbal contents only in the speech addressed to the students of disadvantaged social status. In a frame-by-frame observation of this speech, in fact, we detected a facial expression of contempt, lasting less than a quarter of a second, coherent with his words of indignation for scholastic difficulties solely originated from disadvantaged social conditions that he declared to have personally known. This micro-expression (Ekman and Friesen, 1969), immediately regulated,

revealed Obama's emotion also beyond his communicative intentions. Interestingly, contempt like anger occurs after social or moral transgressions. Nevertheless, unlike anger, contempt only arises when social exclusion of targets is attributed to the unresponsiveness of actors (Fischer and Giner-Sorolla, 2016). Implying a social distancing as well as a self-regulatory function, contempt, therefore, coherently echoed both Obama's severe words against marginalization of low-status students and the immediate regulation of his emotional expression. We decided, therefore, to observe if the presence or the absence of this contempt expression could affect the participants' reactions to this humble Obama's speech.

To answer this research question, we randomly assigned participants either to a condition in which the expression of contempt taken from the Back-to-school speech preceded the reading of the transcript of a part of Obama's speech (experimental condition) or to a control condition in which participants observed a neutral expression of Obama's face before reading the same transcript.

To build up the stimulus for the control condition, we first selected, from the same video, a still image of Obama's face showing, according to FACS analysis (Ekman et al., 1978; Hager et al., 2002), no emotional expressions. We then asked participants to a pilot study to assess the expression shown in both images, choosing from a list of emotions, including "contempt" and "neutral." The results showed that all participants recognized the expression of contempt. However, many participants attributed an emotional value also to the neutral image. We explained these wrong perceptions by the fact that, being caught from a video of Obama speaking, the image had a dynamic quality misleading participants' perceptions. To help participants to recognize Obama's face as a neutral one, we chose, therefore, a posed photo. However, being this photo a well-known one, questions were introduced to control the familiarity of Obama's image and attitudes toward him.

Finally, to avoid the social desirability effects, participants were asked to describe their reactions to Obama's speech following semi-projective instructions of putting themselves in the shoes firstly of a low-status student and then of a high-status student listening to this presidential address.

Hypotheses

Based on the already discussed literature, we expect that participants reading President Obama's self-disclosure of his own school difficulties, after viewing the image showing Obama's contempt, i.e., a facial expression consistent with his words, will perceive his humble speech in more positive terms, evaluating it as more sincere. Instead, we expect that the inconsistency between the verbal content and the neutral face shown in the control condition will provoke an evaluation of falsehood and hypocrisy, leading to a negative perception of the speaker *(Hypothesis 1)*. In addition, we expect a difference in the reactions of participants, depending on identification suggested in semi-projective instructions. Specifically, we expect that identifying with a low-status student will provoke more positive reactions *(Hypothesis 2)*, while identification with a student who does not experience socioeconomic difficulties will cause more negative reactions *(Hypothesis 3)*. In this first explorative study, only aimed at observing effects on receivers of the facial expressions of the speaker, participants' endorsement of gender stereotypes, although taken into consideration of introducing the issue of humble political speech, was neither measured nor controlled.

METHOD

Participants

By means of G*power (Faul et al., 2007), an *a priori* power analysis of the required sample (Lakens, 2022) for ANCOVA was performed with a power of 0.95, an α of 0.05, a medium effect size [f = 0.30, eta square: 0.08, using Cohen (1988) criteria], and one predictor (i.e., familiarity); this analysis indicated that a sample of 152 participants was required. Our sample's size complied with this analysis, since 175 participants (122 women; 52 men; 1 other; average age of 22 years, DS = 4.84; 127 high school graduates and 48 graduates) were recruited as a snowball convenience sample and randomly assigned either to the control (N = 95) or to the experimental condition (N = 80). The study was conducted during COVID-19 restrictions due to online participation. Despite its limitations, we judged the benefit of this first collection of explorative data worthwhile.

Stimuli

As briefly discussed above, the coherence vs. incoherence between facial expression and words was induced using two images of Obama. In the lack of coherence condition (control condition), a well-known photo of Obama's face showing no emotional expression was presented (**Figure 1**). In the coherence condition (experimental condition), we used a still image taken from the video of the "Back-to-school" speech at Wakefield High School (**Figure 2**). Having observed this video through a FACS analysis (Ekman et al., 1978; Hager et al., 2002) in a previous descriptive study (Leone et al., 2018a), a micro-expression of contempt (Left Unilateral Action Unit 14) appeared on his face for less than a quarter of a second when Obama said: *"I get it. I know what that's like*."

The Transcript of the Speech at Wakefield High School

As already discussed, the excerpt chosen as an instance of Obama's humble stance refers to his autobiographical social sharing and runs as follows (you can find the complete video-recording of this "Back-to-school" speech at,[1] being this excerpt running from 9.12 to 9.48 min):

«*I know it's not always easy to do well in school. I know a lot of you have challenges in your lives right now that can make it hard to focus on your schoolwork. I get it. I know what it is like. [.] I was raised by a single mother who struggled at times to pay the bills and wasn't always able to give us things the other kids had. There were times when I missed having a father in my life. I was lonely and felt like I didn't fit in.*

I did some things I'm not proud of and got in more trouble than I should have. And my life could have taken a turn for the worse. But I was lucky».

[1] https://www.youtube.com/watch?v=8ZZ6GrzWkw0

FIGURE 1 | Neutral facial expression accompanying the speech excerpt in the control condition. Image licensed under Creative Common Attribution 3.0 License. Photographed and attributed to Pete Souza. Available via https://it.wikinews.org/wiki/File:Official_portrait_of_Barack_Obama.jpg.

FIGURE 2 | Facial expression of contempt accompanying the speech excerpt in the experimental condition. Screenshot taken from https://obamawhitehouse.archives.gov/copyright. Copyright policy details that pursuant to federal law, government-produced materials appearing on this site are not copyright protected. The United States Government may receive and hold copyrights transferred to it by assignment, bequest, or otherwise. Except where otherwise noted, third-party content on this site is licensed under a Creative Commons Attribution 3.0 License. Visitors to this website agree to grant a non-exclusive, irrevocable, royalty-free license to the rest of the world for their submissions to Whitehouse.gov under the Creative Commons Attribution 3.0 License.

Measures of Dependent Variables

Perception of the Speaker

After reading the text, participants were asked the question: *"Barack Obama's words you just read, how do they look to you?"* To answer the question, participants used a list of 26 adjectives, evaluating each one of them on a scale ranging from 0 (not at all) to 3 (a lot). The adjectives included in the list were either positive (*"affectionate," "authentic," "warm," "understanding," "convincing," "empathetic," "encouraging," "participatory," "reassuring," "sincere," "spontaneous," "touching," and "close")* or negative *("artifact," "exaggerated," "false," "out of place," "cunning," "hypocritical," "manipulative," "paternalistic," "rhetorical," "honeyed," "strategic," "cold," and "weak")*. Adjectives were randomly presented to avoid the order effect.

Emotional Reactions to the Speech

Participants were then asked to put themselves in the shoes of a student in a condition of socioeconomic difficulty listening to Barack Obama's speech and to indicate their reactions to it. Immediately, the same question was proposed, asking participants to identify with a student in medium-high socioeconomic condition. To answer both questions, a list of 32 reactions was presented, asking participants to assess their degree of agreement for each item from 0 (not at all) to 3 (a lot). The reactions were either positive ("*welcomed," "involved," "impressed," "encouraged," "proud," "reassured," "surprised," "inspired," "moved," "empowered," "stimulated," "valued," "understood," and "excited")* or negative *("uncomfortable," "embittered," "bored," "angry," "would have felt contempt," "would have felt shame," "disappointed," "embarrassed," "afraid", "guilty," "indifferent," "indignant," "irritated," "perplexed," "sad," and "offended")* and were randomly presented.

Control Measures

To check familiarity, we asked participants to answer on a scale from 0 (not at all) to 3 (very) to the question "*How familiar is Obama to you?*". To control the opinions held on Obama, we asked participants to answer the question "*Do you think Obama got his position*," choosing among the prearranged answers: *by merit, fortunately, by strategy, or something else*. To check the degree of sympathy toward Obama, we proposed to the participants to evaluate a semantic differential on a 5-step scale from 0 ("*dislike*") to 5 ("*sympathy*") by answering the question *"What do you feel when thinking of Obama?"* Finally, we asked for some personal information, including their political orientation to which participants could respond by positioning themselves in one of the prearranged answers: *Left*, *Center-left*, *Center*, *Center-right*, *Right*, or *something else*.

PROCEDURE

The experimental procedure was carried out through an online questionnaire, created with Google Forms. It allowed us to divide

the questionnaire into several pages, so that the participants could see only one page at a time and cannot go back to previously filled-in pages. Before accessing the questionnaire, participants read the informed consent form and were asked to consent to the processing of data. Subsequently, participants were presented with a first text explaining that every year the President of the United States addresses students with a "Back-to-School" speech and that they will read an excerpt from one of these speeches, delivered by the US President Barack Obama. On the next page, participants were randomly assigned either to the group presented with Obama's photo showing a neutral expression or to the group presented with Obama's image showing an expression of contempt and were free to decide when moving to the next page, where they were presented with the excerpt of Obama's speech. After taking their time for reading the excerpt, participants answered the questions investigating the dependent and control variables. Finally, they gave some information about themselves to better describe the sample.

RESULTS

To examine potential differences between experimental and control conditions (expression of contempt vs. neutral expression) concerning the dependent variables, we run a MANCOVA analysis for each adjective describing both Obama's perception and participants' emotional reactions. To reduce the type I error, we decided to set the significance level at $p \leq 0.025$.

Perception of Obama's Self-Disclosure

The analysis of variance for the adjectives describing the speaker showed the following level of significance, also by taking into account the Obama's familiarity as a covariate: *Sincere* [$F_{(1,175)} = 8.56$, $p = 0.004$, $\eta 2 = 0.047$], *Spontaneous* [$F_{(1,175)} = 6.74$, $p = 0.010$, $\eta 2 = 0.038$], *Touching* [$F_{(1,175)} = 5.47$, $p = 0.020$, $\eta 2 = 0.031$], *Artifact* [$F_{(1,175)} = 8.71$, $p = 0.004$, $\eta 2 = 0.048$], *Exaggerated* [$F_{(1,175)} = 10.59$, $p = 0.001$, $\eta 2 = 0.058$], *False* [$F_{(1,175)} = 9.91$, $p = 0.002$, $\eta 2 = 0.054$], *Cunning* [$F_{(1,175)} = 10.37$, $p = 0.002$, $\eta 2 = 0.057$], *Hypocrite* [$F_{(1,175)} = 8.53$, $p = 0.004$, $\eta 2 = 0.047$], *Rhetoric* [$F_{(1,175)} = 5.75$, $p = 0.018$, $\eta 2 = 0.032$], *Strategic* [$F_{(1,175)} = 12.21$, $p = 0.001$, $\eta 2 = 0.066$], and *Weak* [$F_{(1,175)} = 5.76$, $p = 0.017$, $\eta 2 = 0.032$]. Hence, as shown in **Table 1**, Obama's discourse is perceived as more sincere, more spontaneous, and more touching when it is matched with the image of his expression of contempt. In the control condition, his communication is perceived instead as more artifact, more exaggerated, more false, more cunning, more hypocrite, more rhetoric, and more strategic, but also weaker.

Semi-Projective Self-Assessment of Participants' Reactions

Participants were asked to self-assess their reactions to Obama's autobiographical recall trying to identify themselves firstly with a student in a difficult socioeconomic condition and then with a student in a medium-high economic condition.

Applying a MANCOVA analysis, with Obama's familiarity as a covariate, to self-assessments collected according to the first suggestion (identify with a student of low socioeconomic condition), depending on their experimental conditions, participants differently self-assessed themselves in reference to the adjectives *Interested* [$F_{(1,175)} = 6.08$, $p = 0.015$, $\eta 2 = 0.034$], *Outraged* [$F_{(1,175)} = 5.25$, $p = 0.023$, $\eta 2 = 0.030$], *Perplexed* [$F_{(1,175)} = 10.26$, $p = 0.002$, $\eta 2 = 0.056$], and *Offended* [$F_{(1,175)} = 6.55$, $p = 0.011$, $\eta 2 = 0.037$]. More precisely, participants in the experimental condition declared to be more interested in reading Obama's autobiographical self-disclosure than participants presented with an image showing a neutral expression. Contrariwise, participants declared to be more outraged, more perplexed, and more offended when reading Obama's words after seeing the photo showing a neutral expression (**Table 2**).

Interestingly, participants identifying themselves with a student in a medium-high economic condition, when compared between the experimental and the control condition, did not show any significant difference.

DISCUSSION

Taken all together, the results seem to confirm our first hypothesis. The perception of the same autobiographical social sharing, in fact, deeply varied according to the information conveyed by Obama's facial expression, shown before the reading. We could articulate that coherence between words and facial

TABLE 1 | Average scores and standard deviation of perception's adjectives in the experimental and control conditions (expression of contempt vs. neutral expression).

Perception's adjectives	Expression of contempt		Neutral expression	
	M	DS	*M*	DS
Sincere	2.58	0.63	2.24	0.82
Spontaneous	2.09	0.90	1.73	0.90
Touching	2.09	0.83	1.78	0.88
Artifact	0.25	0.46	0.55	0.78
Exaggerated	0.14	0.35	0.41	0.68
False	0.04	0.19	0.27	0.64
Cunning	0.40	0.70	0.81	0.93
Hypocrite	0.09	0.28	0.34	0.71
Rhetoric	0.56	0.74	0.88	0.97
Strategic	0.80	0.83	1.29	0.99
Weak	0.18	0.47	0.39	0.66

TABLE 2 | Comparison of significantly different medium scores and standard deviation of participants identifying with students in low economic conditions, according to experimental and control conditions (expression of contempt vs. neutral expression).

Reactions	Expression of contempt		Neutral expression	
	M	DS	*M*	DS
Interested	2.39	0.70	2.11	0.78
Outraged	0.06	0.24	0.21	0.52
Perplexed	0.40	0.56	0.74	0.77
Offended	0.06	0.24	0.21	0.46

emotional expression may account for the humble utterance to be sincere, significantly reducing the social distance between this prominent political leader and his audience. On the contrary, the incoherence between his touching words and his neutral face shown before reading could lead perceivers to consider the humble attitude of Obama only as a socially desirable, if not hypocritical, posture. An interesting facet of this data refers to the adjective "weak." Data suggest that a humble speaker appears weak to his audience not when expressing a controversial reaction, such as contempt, but only when his face is not leaking the same indignant attitude conveyed by his words. This could help to better understand why, in previous studies on humble leaders, we found that they were negatively perceived when discussing moral topics (D'Errico, 2019).

Regarding the second hypothesis, only some of our theoretical expectations were proved true. When participants put themselves in the shoes of students of low social condition, they were interested in this unusual communicative move, only when Obama's facial expression was coherent with his words. However, if such passionate and unusual words were associated with a neutral expression, then they backfired, provoking negative reactions instead of social closeness. Interestingly, the falling short of expectancies related to reactions of participants identifying with high-status students suggests that the humble self-exposure of Obama was only fine-tuned for students of low social status: precisely the kind of audience he was addressing during his "Back-to-school" speech. To summarize, the importance of this interconnection between bodily communicative signals, relational humble stance, and potential features of the audience (low or high socioeconomic condition) shed the first light on the complexities of the persuasiveness of the humble political speech. Future studies should better consider other participants' characteristics, such as political engagement and attitude, or individual differences in the social dominance's orientation.

CONCLUSIVE REMARKS

Based on the idea that a humble stance during political speeches may either enhance the positive face of political leaders or backfire against them, this study aimed to explore the consequences of Obama's humble sharing of his school difficulties with students of low social condition, listening to a "Back-to-school" discourse he delivered when he was the US President. In the framework of a multimodal analysis of communication, the Independent Variable manipulated was the coherence vs. the lack of coherence between two facial expressions of Obama (contempt vs. neutral), shown before presenting participants with an excerpt of the talk conveying a brief autobiographical sharing of his school difficulties. Results highlight how much participants' reactions changed, only because of seeing Obama's facial expression of contempt. It could be argued that for receivers, the leaking of a powerful speaker's emotions during an official talk can act as an indirect yet powerful signal of authenticity, especially when an unexpected self-exposure puts a powerful politician to the same level of his audience. However, this is only a first explorative study, presenting some important limitations, and many questions remain unsolved.

The first limitation is represented by the different kinds of images used for our manipulation. In fact, to prevent the distortions of Obama's neutral expression observed in the pilot study, the experimental condition showed a still image taken from Obama's speech, while the control condition showed a posed photo. Moreover, in the original Back-to-school speech, contempt appeared as a micro-expression. Since our procedure aimed at the first exploration of the effects of congruence between facial expressions and verbal contents of a humble political speech, we decided to show this micro-expression as a still image to be observed for several seconds, seriously changing the features of the original case study. Nonetheless, since this first study shows how an emotional expression congruent with words influences participants' perceptions and reactions to a politician's humble speech, we can now design further studies based on new and more sophisticated procedures. The effects of micro-expressions could now be studied as an I.V. *per se,* by importing procedures developed for this field of study (cf. Stewart et al., 2009). For instance, participants could be presented either with the original video, showing the micro-expression, or with a second video, manipulated, where the micro-expression is replaced by a neutral expression, to assure the two videos have the same length. Moreover, the sample size of participants, although complying with an *a priori* power analysis settled at the classic 0.05 level of significance, was slightly under the number requested for the more prudent level of significance applied. Finally, our study did not explore the participants' gender stereotypes, although the literature on humble political speech proved this aspect to be relevant in shaping the participants' expectancies about political speakers. Further studies may change the research's design, making this variable interact both with the multimodal communicative congruence and with different kinds of emotions shown by the speaker.

Many other research questions may be addressed. The effects described in this study could be partially due to the special intimacy produced by the social sharing of autobiographical memories, and other kinds of humble utterances could lead to different reactions. In addition, specificities of contempt must be better explored in comparison with cognate emotions (i.e., rage, sadness). Finally, the effects of humble communication in political speeches different from "monologs" (Bull and Fetzer, 2010) must be analyzed. In fact, humility's consequences could deeply vary, according to the symmetric or asymmetric relations between speakers and audiences. The categorization of political speeches according to a decreasing symmetry (ranging from the highest asymmetry of monologs to the more balanced relationship between politicians and journalists and to the symmetry of Question Times and Parliamentary debates) (Bull and Fetzer, 2010) offers a useful grid to further develop the first ideas presented in this study.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethical Committee of the University of Bari. All the procedures followed the Helsinki ethical principles and ethical codes of AIP (Italian Psychology Association). Informed consent was obtained for all participants. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

AM and GL designed the study. IS and SM realized the procedure and collected the data. All authors analyzed the results and co-wrote the article.

FUNDING

This publication has been partially financed by Ph.D. Course in Psychology and Cognitive Science, Sapienza University of Rome.

REFERENCES

Arendt, H. (1978). *The Jew as Pariah*. New York: Grove Press.

Bull, P., and Fetzer, A. (2010). Face, facework and political discourse. *Rev Int. Psychol. Soc.* 23, 155–185.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*, 2nd Edn. Hillsdale, NJ: Lawrence Erlbaum Associates.

Davis, D. E., Worthington, E. L. Jr., Hook, J. N., and Hill, P. C. (2013). Research on religion/spirituality and forgiveness: a meta-analytic review. *Psycholog. Relig. Spiritual.* 5, 233–241. doi: 10.1037/a0033637

D'Errico, F. (2019). Too humble and sad': the effect of humility and emotional display when a politician talks about a moral issue. *Soc. Sci. Inf.* 58, 660–680. doi: 10.1177/0539018419893564

D'Errico, F., Bull, P., Lamponi, E., and Leone, G. (2022). Humility expression and its effects on moral suasion: an empirical study of Ocasio-Cortez's communication. *Hum. Aff.* 32, 101–117.

D'Errico, F., and Poggi, I. (2019). Tracking a leader's humility and its emotions from body, face and voice. *Web Intelligence* 17, 63–74. doi: 10.3233/WEB-190401.10.3233/WEB-190401

Eagly, A. H. (1987). *Sex Differences In Social Behavior: A Social-Role Analysis*. Hillsdale, NJ: Erlbaum.

Eagly, A. H. (2005). Achieving relational authenticity in leadership: does gender matter? *Leadersh. Q.* 16, 459–474. doi: 10.1016/j.leaqua.2005.03.007

Ekman, P., and Friesen, W. V. (1969). Nonverbal leakage and clues to deception. *Psychiatry* 32, 88–105. doi: 10.1080/00332747.1969.11023575

Ekman, P., Friesen, W. V., and Hager, J. (1978). *The Facial Action Coding System (FACS): A Technique for the Measurement of Facial Action*. Palo Alto, CA: Consulting Psychologists Press.

Faul, F., Erdfelder, E., Lang, A. G., and Buchner, A. (2007). G* Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav. Res. Methods* 39, 175–191. doi: 10.3758/bf03193146

Fetzer, A., and Bull, P. (2012). Doing leadership in political speech: semantic processes and pragmatic inferences. *Discourse Soc.* 23, 127–144.

Fischer, A., and Giner-Sorolla, R. (2016). Contempt: derogating others while keeping Calm. *Emot. Rev.* 8, 346–357.

Gunn, E. (2015). *Mediated Authenticity: How the Media Constructs Reality*. New York: Peter Lang.

Hager, J. C., Ekman, P., and Friesen, W. V. (2002). *Facial Action Coding System*. Salt Lake City, UT: A Human Face.

Lakens, D. (2022). *Sample Size Justification*. Collabra: Psychology.

Leone, G., Antenore, M., and Serlupi Crescenzi, L. (2018a). Condivisione ed empowerment: l'uso dei ricordi autobiografici in due discorsi di Barack Obama a studenti di alta o bassa classe sociale [sharing and empowerment: the use of autobiographical memories in two speeches by Barack Obama to students of high or low social class]. *G. Ital. Psicol.* 45, 217–244.

Leone, G., Giner-Sorolla, R., D'Errico, F., Migliorisi, S., and Sessa, I. (2018b). It's time to be ashamed! Reactions to the breaking of a long-lasting self-censorship on ingroup war crimes. *Test. Psychom. Methodol. Appl. Psychol.* 25, 519–535. doi: 10.4473/TPM25.4.4

Leone, G., Di Murro, F., and Serlupi Crescenzi, L. (2015). "From personalization to parrhesia: a multimodal analysis of autobiographical recalls in Barack Obama's political speech," in *Conflict and Multimodal Communication*, eds F. D'Errico, I. Poggi, L. Vincze, and A. Vinciarelli (Cham: Springer), 349–374.

Liu, C. (2016). Does humble leadership behavior promote employees' voice behavior? A dual mediating model. *Open J. Bus. Manag.* 4, 731–740.

Luebke, S. M. (2020). Political authenticity: conceptualization of a popular term. *Int. J. Press. Politics* 26, 635–653. doi: 10.1177/1940161220948013

Scardigno, R., Papapicco, C., Luccarelli, V., Zagaria, A. E., Mininni, G., and D'Errico, F. (2021). The humble charisma of a white-dressed man in a desert place: pope Francis' communicative style in the Covid-19 pandemic. *Front. Psychol.* 12:683259. doi: 10.3389/fpsyg.2021.683259

Stewart, P. A., Waller, B. M., and Schubert, J. N. (2009). Presidential speech making style: emotional response to micro-expressions of facial affect. *Motiv. Emot.* 33, 125–135.

Tangney, J. P. (2000). Humility: theoretical perspectives, empirical findings and directions for future research. *J. Soc. Clin. Psychol.* 19, 70–82. doi: 10.1521/jscp.2000.19.1.70

Weidman, A. C., Cheng, J. T., and Tracy, J. L. (2018). The psychological structure of humility. *J. Pers. Soc. Psychol.* 114, 153–178. doi: 10.1037/pspp0000112

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Mastropietro, Bull, D'Errico, Sessa, Migliorisi and Leone. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers | Frontiers in Communication

TYPE Original Research
PUBLISHED 04 August 2022
DOI 10.3389/fcomm.2022.909913

Check for updates

OPEN ACCESS

EDITED BY
Oliver Niebuhr,
University of Southern Denmark, Denmark

REVIEWED BY
Dani Byrd,
University of Southern California, United States
Margaret Zellers,
University of Kiel, Germany

*CORRESPONDENCE
Tamara Rathcke
tamara.rathcke@uni-konstanz.de

SPECIALTY SECTION
This article was submitted to Language Sciences, a section of the journal Frontiers in Communication

RECEIVED 31 March 2022
ACCEPTED 18 July 2022
PUBLISHED 04 August 2022

CITATION
Rathcke T and Fuchs S (2022) Laugh is in the air: An exploratory analysis of laughter during speed dating. *Front. Commun.* 7:909913. doi: 10.3389/fcomm.2022.909913

COPYRIGHT
© 2022 Rathcke and Fuchs. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Laugh is in the air: An exploratory analysis of laughter during speed dating

Tamara Rathcke[1,2]* and Susanne Fuchs[3]

[1]University of Konstanz, Konstanz, Germany, [2]University of Kent, Canterbury, United Kingdom, [3]Leibniz Zentrum Allgemeine Sprachwissenschaft, Berlin, Germany

Laughter is a ubiquitous vocal behavior and plays an important role in social bonding, though little is known if it can also communicate romantic attraction. The present study addresses this question by investigating spontaneous laughter produced during a 5-min conversation in a heterosexual speed-dating experiment. Building on the posits of Accommodation Theory, romantic attraction was hypothesized to coincide with a larger number of shared laughs as a form of convergence in vocal behavior that reduces the perceived distance between the daters. Moreover, high-attraction dates were expected to converge toward the same laughter type. The results of the experiment demonstrate that (a) laughs are particularly frequent in the first minute of the conversation, (b) daters who are mutually attracted show a significantly larger degree of temporal overlap in laughs, (c) specific laughter types (classified as a nasal "laugh-snort") prevail in high-attraction dates, though shared laughs are not consistently of the same type. Based on this exploratory analysis (limited to cisgender, heterosexual couples), we conclude that laughter is a frequent phenomenon in speed dating and gives some indication of a mutual romantic attraction.

KEYWORDS

vocal accommodation, laughter, speed dating, attraction, timing, respiration

Introduction

Laughter is a universal communicative behavior that is observed and recognized in all cultures around the world (Lefcourt, 2000; Provine, 2000; Kipper and Todt, 2001; Sauter et al., 2010; Bryant et al., 2016). Despite its prevalence in social interaction (e.g., Jefferson, 2004; Vettin and Todt, 2004; Tanaka and Campbell, 2011; Truong and Trouvain, 2014), social meanings and communicative functions of laughter are still not entirely understood. In a large cross-cultural study, Bryant et al. (2016) investigated the perception of laughter produced in conversations between friends or strangers and showed that listeners from 24 different cultures could reliably identify the familiarity levels between speakers based on their laughs alone. This finding suggests that laughter might fulfill an important social function of signaling in-group/out-group relationships between conversational interlocutors. It has further been proposed that laughter might have a face-preserving function in situations of a conversational face threat (Grammer and Eibl-Eibsfeldt's, 1990). Specifically, "mixed sex encounter have a high potential risk

of face loss and rejection [...]. A statement accompanied by laughter which has the metacommunicative function of signaling that this statement is a play statement could reduce this possibility" (Grammer and Eibl-Eibsfeldt's, 1990, p. 195-196). Laughter has also been suggested to reflect and reinforce positive emotion in social interaction (Owren and Bachorowski, 2003; Smoski and Bachorowski, 2003). Both in-group/out-group relations, face preservation and positive emotion might play an important role in the context of an emerging romantic attraction. The present paper aims to address the empirical question of how, and to what extent, laughter is involved in mutual romantic attraction, by studying spontaneous laughter produced in the context of cisgender, heterosexual speed dating.

Acoustic and respiratory features of laughter

Bryant et al. (2016) showed that laughter can be acoustically diverse and analyzed a number of laughs that had been previously investigated in a perception study. Accordingly, laughter produced among friends has a shorter duration and larger variability in the fundamental frequency and intensity, compared to laughter produced among strangers.

In an earlier large-scale study, Bachorowski et al. (2001) analyzed acoustic features of 1,024 laughs produced by 97 young adults while watching humorous video clips. The study dismantled the myth of a "stereotypical laugh"—a harmonically rich, vowel-like vocalization consisting of multiple repetitions of "ha," "he" or "ho." Merely 30% of all observed laughs were of this type. In contrast, around 50% of all recorded laughs in this study were unvoiced, with two sources of voiceless turbulence in the signal differentiating between a nasal, "snort-like" and an oral, "grunt-like" laugh. Among acoustic parameters that play a role in laughter, fundamental frequency, signal-to-noise ratio, glottal excitation features, vowel quality, intensity and duration are often mentioned (Grammer and Eibl-Eibsfeldt's, 1990; Bachorowski et al., 2001; Kohler, 2008; Tanaka and Campbell, 2011), though there are disagreements as to how these acoustic characteristics should be mapped onto the auditory impressions of laughter and lead to an acoustically informed laughter typology (Trouvain, 2003).

Physiological correlates of laughter can be found in a forced exhalation (Ruch and Ekman, 2001), sometimes described as the "spasms of the diaphragm" (Scott et al., 2014), and in repeated contractions of the intercostal muscles within the chest wall (Kohler, 2008). A strong physiological component that plays a role in laughter is respiration, though only a very limited number of studies has addressed the respiratory kinematics of different laughs. Filippelli et al. (2001) studied respiration in participants watching a variety of humorous video clips and found a steep decrease in lung volume during laughter (to a level as low as the residual respiratory capacity in some individuals). However, such respiratory changes might depend on the type of laughter which was not controlled for in Filippelli et al. (2001). For example, a rapid drop in lung volume requires the glottis to be open, meaning that the vocal fold vibration that leads to voicing cannot take place. This appears to be at odds both with the voiced hahaha-type of laughter reported in the literature (e.g., Provine, 1996, 2000; Bachorowski et al., 2001; Trouvain, 2003) and with the frequent co-occurrence of speech and laughter (Nwokah et al., 1999). A physiologically imposed limit applies to the frequency of pulses within a laugh cycle. These can vary from 4 to 12 pulses per cycle, and their occurrence is influenced by the individual lung volume (Ruch and Ekman, 2001).

Respiratory kinematics of laughter during spoken social interaction has rarely been addressed in previous research. McFarland (2001) investigated interpersonal synchrony of breathing kinematics using cross-correlation of data taken from spontaneous conversations between same-sex participants and reported many occurrences of simultaneous laughter close to turn changes. Interlocutors showed a high degree of joint respiratory synchronization (in-phase coordination of in-breaths) around 5 s before or after a turn change, or after a shared laugh. Shared laughter refers to those periods in a conversation when both interlocutors are laughing simultaneously. Even though respiratory kinematics involved in single or shared laughter is rather poorly documented, mutual influences between respiration and basic emotions have been frequently discussed, with laughter being often seen as part of the basic emotion repertoire (specifically in the category of joy). Empirical evidence for the interplay between laughter, emotion and breathing dates back to the work by Feleky (1915) who found that basic emotions—among them joy with laughter—have specific breathing characteristics and may express themselves in the facial and the respiratory muscles. Bloch et al. (1991) elaborated that joy-laughter has some specific temporal relations between inhalations and exhalations, namely the ratio between the duration of an inhalation and an exhalation was ∼0.4. This timing ratio was further accompanied by a small inhalation amplitude and saccadic patterns in the exhalation amplitude.

Frequency of individual and shared laughter in conversation

As far as an overall frequency of laughter in a conversation is concerned, about 5 laughs per 10 min have been observed in daily interactions among close friends (Vettin and Todt, 2004), though laughter frequency is known to increase up to 75 times per 10 min in the context of an interaction between strangers of opposite sex, with women laughing slightly more frequently than men (Grammer, 1990). The most frequently observed laughs are known to be purely interactional, unrelated

to the presence of humor (Jefferson, 2004; Vettin and Todt, 2004). These interactional laughs are often initiated by the person who has just spoken (Glenn, 1989, 1992; Vettin and Todt, 2004), occur at the end of utterances (Provine and Emmorey, 2006) and are said to express emotion of the speaker (; Provine, 2004; Provine and Emmorey, 2006; Kohler, 2008). A prolonged period of laughter in a conversation offers an opportunity for the interlocutor to join in, and share the laugh (Glenn, 1989, 1992, following Jefferson, 1979, p. 3). The sequential order of shared laughs is sometimes described as an 'accepted invitation to laugh'. Such joint laughter has been identified as a cross-culturally universal means of signaling affiliation and expressing social cohesion (Bryant et al., 2016). Shared laughter has also been identified as a behavioral indicator of wellbeing in romantic relationships where closeness and social support are positively associated with the frequency of shared laughs (Kurtz and Algoe, 2015).

Functions of laughter

Our present understanding of the potential role and functions of laughter in romantic communicative interaction is limited to cisgender, heterosexual dyads. A recent meta-analysis by Montoya et al. (2018) suggests that the social importance of laughter extends as far as fostering romantic bonds and expressing sexual attraction, though empirical evidence in support of this view is not always clear-cut. For example, a study by Grammer (1990) found no correlation between the frequency of laughter and a romantic interest in conversations of heterosexual interlocutors of the opposite sex. According to the results of the study, laughter can signal either aversion or interest. It is rather the body posture that clearly indicates romantic intentions (Grammer, 1990). In contrast, McFarland et al. (2013) study showed that heterosexual men, but not women, use laughter as an indicator of their attraction. This conflicting evidence might be due to cross-cultural differences in mating behavior. While Grammer and Eibl-Eibsfeldt (1990) sample was German, McFarland et al.'s (2013) conducted their study with American participants. A discrepancy in the results is also very likely to arise from the difference in the methods of the data collection employed by the two studies, given that specifics of the interactional context and the surrounding environment can affect linguistic behavior (e.g., Hay and Drager, 2010; Hay et al., 2017). While Grammer (1990) did not inform the study participants that a potential romantic interest arising during their interaction was the purpose of the experiment, McFarland et al.'s (2013) set out to collect their data in an explicit speed dating setting.

Speed dating has been recognized as an ecologically valid means of studying the role of interpersonal attraction on linguistic behavior (McFarland et al., 2013; Michalsky et al., 2017) and is a generally well-known method of studying mate selection in social sciences (e.g., Finkel et al., 2007). A speed dating event usually attracts singles who seek to meet many strangers within a short period of time, with a view to find a romantic partner (Finkel and Eastwick, 2008; McFarland et al., 2013). Each date lasts just 2–3 min and offers two strangers a short window of opportunity to find out if the interlocutor has the potential of becoming a romantic partner. When the time is up, a sound signal indicates that participants have to move on to a new dating partner. As McFarland et al.'s (2013, p. 1,605) note, speed dates share many characteristics with initial romantic conversations in other contexts: "people meet and greet one another, they try to reveal positive features of themselves and learn about the other, they engage in efforts to relate and connect with one another, and they experience (a)symmetries of attraction." A dater might meet up to 20 mating candidates at such an event. After the event, a scoring card is filled in and returned to the organizers who ensure that only those daters who both found each other attractive and expressed an interest in seeing each other again, receive each other's contact details.

Existing studies into the role of laughter in romantic attraction have rarely investigated the role of different laughter types (in fact, it is often unclear if only the vowel-like, stereotypical hahaha-laughter was included in the analyses). However, research by Grammer and Eibl-Eibsfeldt's (1990) indicates that different types of laughter may play a differential role in promoting attraction. In an experiment, Grammer and Eibl-Eibsfeldt's (1990) observed dating behaviors of cisgender, heterosexual couples and concluded that male participants tended to express more interest in seeing a female stranger again if she produced voiced, but not if she produced unvoiced, laughter during their brief interaction.

Aims and hypotheses of the study

By adopting a more nuanced, empirically elaborated laughter typology, the present study aims to better understand how laughter is used and produced in the context of an emerging romantic attraction during speed dating. It further aims to make the following novel contributions to the existing field:

(1) It will clarify what types of laughter prevail in high vs. low levels of mutual attraction, and supply empirical evidence on temporal and distributional features of these laughter types;
(2) It will provide examinations of breathing kinematics in different laughter types and evaluate timing, amplitude and the degree of the respiratory overlap in shared laughter during speed dating;
(3) It will assess whether or not shared laughter is moderated by mutual attraction among the daters.

We hypothesized that given its apparent importance in fostering social bonds (Montoya et al., 2018) and improving relationship quality (Kurtz and Algoe, 2015), laughter would play an important role during speed dating and might help to express romantic attraction. Sonorous laughter rich in vocalic resonance was assumed to be particularly likely to accompany high-attraction dates (cf. Grammer and Eibl-Eibsfeldt's, 1990). Moreover, attraction might be signaled *via* an increased number of shared laughs (cf. Kurtz and Algoe, 2015; Bryant et al., 2016). Building on the posits of the Accommodation Theory (Giles, 1973; Giles et al., 1991), we predict to find evidence for laughter convergence (cf. Ludusan and Wagner, 2022). Specifically, shared laughter and laughter of the same type is expected to prevail among daters with high but not low levels of mutual attraction. That is, we predicted a high degree of laughter convergence to express attraction. Giles and Ogay (2007) specifically hypothesize that situations with a high romantic potential might show gendered accommodation patterns. With regards to acoustic features and breathing kinematics, we expected different laughter types to be consistently distinguished by duration and exhalation gradients. Synchronization of the physiological signals between the daters could be a further indicator of mutual attraction (cf. McFarland, 2001).

Methodology

Participants

All participants were registered on a database at the Leibniz-Center for General Linguistics in Berlin and recruited in October 2017. The call for participation invited heterosexual, single males and females, native speakers of German aged between 20 and 30 years, with an interest in speed dating. Two males (henceforth m1 and m2) and six females (henceforth f1-6) volunteered to take part. They were informed about the purpose of the study, signed a consent form, and received a standard rate of €10/h as compensation for their time involvement. Prior to the experiment, the participants had to fill in an online questionnaire and provide some detailed information about their relationship status, previous dating experience and answer questions from a range of socio-psychological scales.

According to the questionnaire, the participants had been single for variable periods of time, ranging from just 2 months (m1) to their whole adult life (m2 and f2). Apart from the females f4 and f6, all participants had some previous experience with online dating, but none of the participants had ever tried speed dating prior to the experiment.

Experimental setup

To bring some romantic mood into the lab, the recording room was decorated with green plants, posters, and flowers (cf. Hay and Drager, 2010, see Figure 1). All computers were moved outside of the lab, without disrupting the set-up that was necessary for tracking multi-channel recordings in real time.

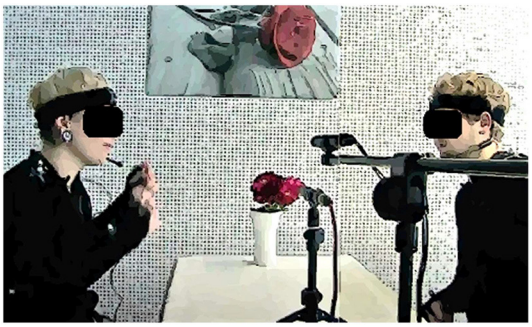


FIGURE 1
Laboratory setting during the experiments showing participants during a conversation (both participants were wearing head-mounted microphones, motion capture jackets and headbands, the respiratory belt was worn underneath the jacket).

All recordings were made with a multimedia setup. To record upper body movement (including head and arms), a motion capture system (OptiTrack, Motive Version 1.9.0) with 12 cameras (Prime 13) was used. An inductance plethysmograph enabled a simultaneous recording of respiratory movements of the ribcage and the abdomen using two belts. Head-mounted microphones and two free-standing microphones were used in parallel to support the temporal synchronization of the signals collected across the different systems. The acoustic stereo signals of the dyadic conversations were recorded *via* two head-mounted microphones (see Figure 1) and stored as a stereo signal on a digital audio tape (DAT) recorder. One of the free-standing microphones was linked to the motion capture data *via* optitrack, and the other free-standing microphone was coupled with the inductance plethysmography signals on a multi-channel recorder. At the beginning and end of each experimental session, synchronization impulses were sent from the motion capture system to the computer connected to the plethysmograph *via* an OptiTrack synch box, which made a subsequent signal synchronization possible. Acoustic data from the three channels (OptiTrack, Plethysmograph and DAT) were also used to check the temporal synchronization of all recordings using cross-correlation.

Procedure

The participants were asked to come to the laboratory on two consecutive days. On the first day, they had a chance to familiarize themselves with the lab, the experimental setting and the equipment. They could ask any questions about the procedure. Their motion-capture jackets and respiratory

belts were fitted by a trained researcher. Baseline recordings were obtained, including a 5-min dialogue with a same-sex conversation partner (i.e., a male or a female confederate). The participants practiced the use of their smartphones for the collection of attraction ratings. At the beginning and the end of the baseline recording, each participant was asked to rate the confederate on a 10-point Likert scale, stating how attractive they found the same-sex confederate (from 1, being the least attractive, up to 10, being the most attractive). In addition, participants performed a series of chest movements in quiet breathing and in different respiratory maneuvers (iso-volume, vital capacity, see Hixon et al., 1973). The iso-volume maneuvers were needed to determine a constant in the habitual movements of the rib cage and the abdomen during quiet breathing. Using the constant derived from quiet breathing, rib cage and abdomen signals could be summed up to derive a joint signal representative of a proportion of the overall lung volume. The vital capacity maneuver served as a reference for each speaker's individual lung volume. For each laughter, its duration (in ms), displacement (in Volt) and the first velocity peak (in Volt/ms) were calculated from the respiratory data and used as dependent variables in the statistical modeling below.

The speed dating experiment took place on the following day. Both male speakers talked to all six female participants, which resulted in 12 dyadic conversations in total. We ensured that participants would not see each other before the recording and directed them to a waiting room upon arrival (there was a separate waiting room for male and female participants). While waiting, participants prepared themselves for the recording session ahead by putting on a motion-capture headband and jacket with markers. Two assistants were present to support with the set-up and the change-over.

During each dating session, participants were first invited to sit down at a table. Their respiratory belts and head-mounted microphones were then connected to the recording computers *via* a set of leads. The experimenters checked all signals and left the recording room as quickly as possible, to give the daters a feeling of privacy. Using a smartphone, the daters rated each partner on a 10-point Likert scale (from 1, being the least attractive, up to 10, being the most attractive) before and after their 5-min dating conversation. An experimenter ensured that the initial attraction ratings were given as soon as the two daters sat at the table and prior to any conversation between the daters taking place. The same experimenter ended each dating session after the allocated 5 min, by entering the room and thanking the participants. It was agreed that the contact details of the daters would be communicated by email, if (and only if) there was a match. That is, an expression of further interest in a follow-up date that was mutual. If only one of the daters expressed further interest in a follow-up date with the interlocutor, no contact details were released. Expressions of further interest were collected on smartphones at the end of each experimental session, right after the final attraction scores were obtained.

The experiment yielded one perfect match: m1 and f2 were both interested in seeing each other again (see Table A3). The romantic interest was not mutual in several pairs: f1 and f3 were interested in a second date with m1 (but not vice versa). In contrast, both male participants expressed an interest in another date with f2, f4, f5, f6, but not f1 or f3. M2 was less popular with all female participants.

The above chemistry (or the lack thereof) was also reflected in the attraction scores collected before and after the speed date (see Tables A1, A2). M1 received higher first-impression scores than m2 (mean scores before the date: 4.2 vs. 2.3, respectively), and all females increased their attraction scores by 1–3 after their date with m1 (mean score after the date: 6.3). In contrast, females felt only slightly (if at all) more attracted to m2 after the date (mean score after the date: 3.2). The wilcoxon signed-rank test showed significantly lower scores for m2 than m1 after the date ($V = 21$, $p < 0.05$), though the attraction-score difference prior to the date did not reach significance ($V = 10$, $p = 0.098$). Both men used the higher end of the given 10-point attraction scale when expressing their attraction to women: scores given by m1 averaged around 7.2. Similarly, scores given by m2 were around 7.0, with the lowest score of 4 (given to f1 before/after the date) and the highest score of 9 (given to f2, f4, f5, f6 after the date). Overall, both males gave significantly higher attraction scores to all women after their date, as compared to their first impression prior to the date ($v = 4$, $p < 0.5$), though the average numerical score difference was rather small (7.4 vs. 6.7) and did not hold for f3 who received a slightly lower score from m1 after their date. These results suggest that the interaction during the date had an impact on the attraction among the daters.

Using the after-date attraction scores and the expressed interest in a second date (0 = both partners uninterested, 1 = only one partner interested, 2 = mutually interested partners), we calculated a cumulative mutual attraction score for each dialogue. This derived score will serve as a predictor in the statistical analyses below. Individual attraction and mutual scores can be found in Tables A1–A3.

Data preparation

All dialogues were first transcribed in Praat (version 6.0.37, Boersma and Weenink, 2018), using the audio signals recorded by the plethysmograph. Transcribers (two native speakers of German) were also instructed to identify bouts of laughter for each speaker, and to annotate those on a separate, "non-verbal" tier. The laughs ter annotations were then inspected in more detail and corrected where necessary, using the stereo acoustic signal from the DAT recorder. Unfortunately, the recorder had a temporary failure that affected conversations of m2 with f1 and f2, resulting in the lack of stereo recordings for the m1_f1 and m1_f2 pairs. Accordingly, all annotations for these pairs were conducted on the basis of the mono-signal from the

respiratory system and are therefore less precise. To account for this issue, we added recording source as a random effect to our statistical modeling.

Following Bachorowski et al. (2001), we adopted a three-way classification of laughs. Accordingly, voiced laughter with a high vocalic proportion was identified as "song-like." Predominantly voiceless laughter was examined with respect to its perceptually salient noise, and classified either as "snort-like" (if the turbulence appeared to originate from the nasal cavity) or "grunt-like" (if the turbulence was produced in the oral or laryngeal tract, including breathy laughs and harsher cackles). As in Bachorowski et al. (2001), an audible inhalation or exhalation noise abutting a laugh episode was not included into the laughter interval. Given that our study, unlike the study by Bachorowski et al. (2001), involved spoken interactions, we also had to add "laughed speech" (or speech-laughs, Nwokah et al., 1999) as the fourth category to account for all observed occurrences of laughter. Following Nwokah et al. (1999), speech-laugh was identified in all cases of a simultaneous occurrence of speech and laughter. We avoided the temporal separation of a long laugh into several intervals, unless there was a prolonged (i.e., longer than 600 ms, cf. Jaffe and Feldstein, 1970) silent pause between two consecutive bouts.

A preliminary annotation was prepared by a trained junior phonetician and subsequently checked by both co-authors. No formal cross-rater agreement was collected. Instead, each laugh was checked by at least two raters and had to be agreed by all raters in less clear-cut cases. Some laughs were difficult to assign to one category as the quality of the laugh might have changed throughout the interval. In (rare, $n < 10$) cases like this, we applied the preponderance principle and classified the laugh based on the temporally and auditorily predominant vocalization (see Figure 2 for examples).

Timing of the respiratory signals of the speakers was identified with reference to the synchronization impulses sent from the motion capture PC to the PC running the Inductance devices and cut using MATLAB (R2017b). After combining the ribcage and abdominal data into one signal representative of overall lung volume, the data were low-pass filtered with a cut-off frequency of 10 kHz and the 6th-order Butterworth filter. The filtered respiratory signal was then used to calculate the velocity as the first derivative.

The auditory annotations of laughter described above served as the points of reference for the examination of the accompanying respiratory curves. An exhalation phase coinciding with an acoustic duration of a laugh was identified as its respiratory signature. Its onset was defined as the point of a zero crossing prior to a stark decrease in the signal velocity. Its offset was more difficult to determine visually because the exhalation gradient was often shallow toward the end of an exhalation cycle, with zero crossings occurring rather frequently. If we took a simple numeric threshold criterion of 10 or 20% to define the on- and offset of the breathing kinematics (as it is common in articulatory analyses, e.g., Katsika et al., 2014), the algorithm would stop after the first bout. However, many laughs consisted of several bouts, each comprising a drop in displacement and a relatively stable part. We therefore considered the last zero crossing before the onset of the next inhalation to constitute the most consistent, reliable, and replicable point for the annotation of the kinematic offset of each laugh.

An example extracted from the dyad of m1 and f4 is shown in Figure 3. The two upper panels display the acoustic waveforms of the vocalizations produced by the two speakers. Laughter duration is indicated by the vertical dashed/dotted lines. Shared laughter starts with the onset of the laugh by m1 and ends with the offset of the laugh by f4. The two lower panels display the accompanying respiratory kinematics. The onset and the offset of the respiratory movements of the two laughs are also indicated by the vertical dashed/dotted lines. A visual comparison of the upper and the lower panels is indicative of the magnitude of potential discrepancies between the acoustics and the kinematics of laughter. In the given example, the acoustic and respiratory duration are tightly aligned in f4 but not in m1 whose respiratory maneuvers start before the acoustic onset of the laugh and end a few milliseconds after. Across all dyads, the duration of shared laughter identified in the acoustic waveforms ranged between 5 and 1,890 ms. The duration of the accompanying respiratory overlaps ranged between 26 and 1,492 ms.

Differences in the respiratory volume between the onset and offset of each laugh were normalized and expressed as percent of the speaker's vital lung capacity. Moreover, we measured peak velocity of the initial respiratory drop (if several velocity peaks occurred, only the first one was considered) and identified laughter duration preceding or following a shared bout (i.e., time lags between the two onsets and the two offsets of a shared laugh, cf. Figure 3)[1].

Data analyses

All statistical analyses reported here were carried out in RStudio (running R version 3.6.1) using the packages lme4 (Bates et al., 2015), lmerTest (Kuznetsova et al., 2017) and ordinal (Christensen, 2018).

1 Since internal clocks of the DAT and plethysmograph recorders showed slight discrepancies, subtle temporal misalignments between the two types of annotations (up to 35 ms per 5 min of speech) could not have been avoided. Although the misalignment did not have any implications for the comparability of the temporal measurements derived from these annotations, the results below describe either the acoustic or the respiratory channel, without directly comparing temporal relationships between the two channels.

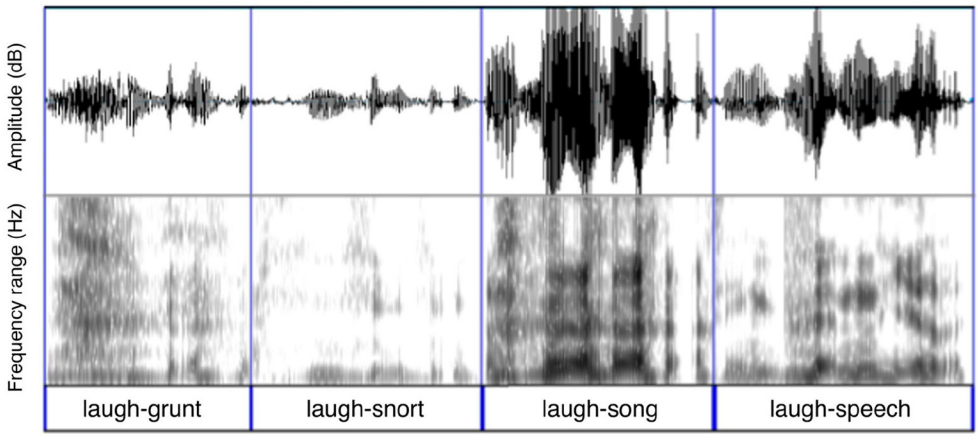


FIGURE 2
Waveforms **(Top)** and spectrograms **(Center)** and annotations **(Bottom)** of the four laughter types produced by the male speaker m1 on his date with f1.

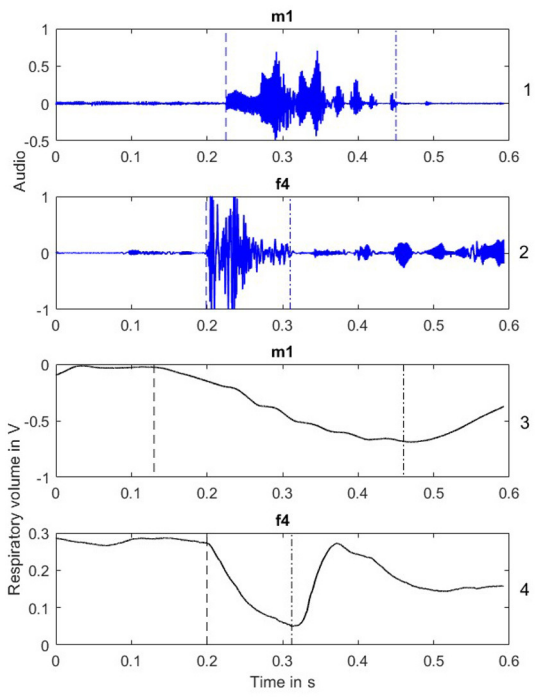


FIGURE 3
Waveforms of a laugh produced by m1 (first panel, see right y-axis for panel number) and f4 (second panel). The accompanying respiratory kinematics (respiratory volume in V, arbitrary unit) are displayed in panel 3 (m1) and panel 4 (f4). Vertical dashed lines mark the onsets of laughter and dashed-dotted lines the offsets.

Models were based on a mixed-effects, statistics with an alpha level of 0.05. Ordinal (for scales), Poisson (for counts) and linear (for duration, displacement and velocity) models were fitted as appropriate. In all mixed models, the random effect structure included intercepts only since models with the maximal structure frequently showed model convergence issues (though whenever possible, we also ran models including speaker-specific random slopes and verified all effects reported below). Dater pair, serial order of a dialogue during the day-long session, and where appropriate, speaker were defined as random effects. If the residuals of the model were not linearly distributed, the data were log-transformed (duration, displacement). Attraction scores (see Section: Distribution of shared and individual laughter), laughter types (see Section: Respiratory properties of shared and individual laughter) and the presence of laughter overlap (see Sections: Time course and duration of laughter, Respiratory properties of shared and individual laughter, and Is there a good predictor for growing attraction?) were fitted as fixed predictors in a forward-fitting procedure (i.e., starting with a simpler model and successively increasing its complexity by adding more factors).

Results

Distribution of shared and individual laughter

An overview of the distribution of all individual and shared laughs is given in Table 1. These count data were analyzed using the Wilcoxon signed rank test for paired samples. Overall, female and male speakers did not significantly differ in the frequency of their laughter (19–22 times in 5 min on average, Wilcoxon signed rank test n.s.).

Shared laughter (defined as a temporal overlap in the acoustically salient laughter bouts, see Figure 3) was also influenced by attraction, with more overlaps found in pairs who

TABLE 1 An overview of the distribution of laughter across the 12 dating pairs (m, male; f, female).

Pair	Attraction score	Laughs (total)	Laughs (m)	Laughs (f)	Shared (%)
m1_f1	4.7	46	18	28	57
m1_f2	5.7	37	18	19	54
m1_f3	4	39	24	15	46
m1_f4	5.7	54	23	31	63
m1_f5	5.3	37	19	18	43
m1_f6	4.7	32	19	13	31
m2_f1	2.3	39	14	25	15
m2_f2	4.3	26	12	14	23
m2_f3	2.7	24	15	9	33
m2_f4	4	58	21	37	31
m2_f5	5	69	31	38	34
m2_f6	4	36	15	21	28
Mean (SD)	4.4 (1.1)	41.5 (13.4)	19.1 (5.2)	22.3 (9.5)	38.2 (14.5)

showed a higher degree of mutual attraction as predicted ($z = 3.3$, $p < 0.001$). Men led these shared laughs more often than women did (65.3% of all cases, $V = 70$, $p < 0.05$). In other words, women joined in a laugh initiated by their interlocutor more frequently than men (who laughed more if they experienced a sense of attraction).

In terms of the four laughter types (see Section: Procedure), the most frequent one was a grunt (34%) while the least frequent was a snort laugh (21%). The difference in this frequency of occurrence was significant ($V = 65.5$, $p < 0.05$). Being a rare laugh, a snort laugh proved the only type of laughter whose occurrence was moderated by the presence of attraction between the conversation partners ($z = 4.3$, $p < 0.001$). A snort-like laugh was more likely to occur when the conversation partners felt attracted to each other. A song-like laugh initiated a shared bout significantly more often (34%) than a speech-laugh did (19%, $V = 66.5$, $p < 0.05$). However, it did not significantly differ from a grunt-like laugh (21%, $V = 45$, n.s.) or snort-like laugh (26%, $V = 32$, n.s.) in this regard. Only 25% of all overlapped laughs were of the same laughter type, and they were not influenced by the mutual attraction among the daters. No durational differences were found in shared laughs initiated by the different laughter types.

Time course and duration of laughter

As shown in Figure 4, laughter prevailed at the very beginning of most dates and declined in frequency over the time course of 5 min. 26.7% of all laughs recorded in this study were produced within the first minute of the conversation (vs. merely 1.6% in the last minute). Wilcoxon signed rank test with continuity correction showed that the decrease in laughter frequency after the first minute of the date was significant (comparing frequencies recorded within the first minute with the average frequencies recorded later in the conversation, $V = 74.5$, $p < 0.01$). Frequent laughter at the start of a conversational encounter is unusual in same-sex dialogues (e.g., no more than 2.3% of all recorded laughter bouts reported in Vettin and Todt, 2004) and might be an important feature in the context of dating (cf. Grammer and Eibl-Eibsfeldt's, 1990).

The acoustic duration of laughter did not differ between male and female speakers and was unaffected by their attraction. Instead, it varied systematically across the different laughter types [$F_{(3)} = 25.6$, $p < 0.001$] and in the presence of an overlap [$F_{(1)} = 27.7$, $p < 0.001$].

Raw acoustic durations of the four laughter types are plotted in Figure 5 and show that the grunt-like laugh (~430 ms on average) was 200–250 ms shorter than all other laughter types (all comparisons significant at $t > 6.0$, $p < 0.001$). The song-like laugh was the longest (~690 ms on average) but did not differ significantly from either speech-laugh or from the snort-like laugh. When laughter was shared, it was about 215 ms longer than individual laughs by the speakers ($t = 5.4$, $p < 0.001$). The lengthening effect of shared laughter applied equally across all laughter types. The duration of the overlap itself was not influenced by the variables of interest, and neither were the duration of the onset or the offset of a shared laugh (see Figure 3). By and large, the incoming interlocutor joined into an ongoing laugh ~320 ms after its start. Once an interlocutor stopped laughing, their partner continued to laugh on their own for further 360 ms on average. This pattern did not show an effect of the interlocutor gender or the degree of mutual attraction.

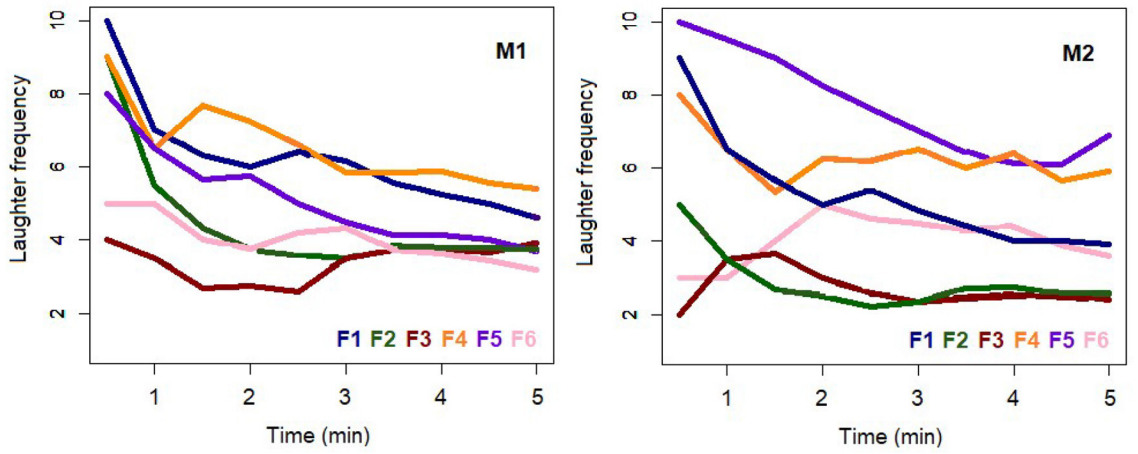


FIGURE 4
Time course of laughter frequency during the twelve dates. Data of the six female speakers are color-coded, data of speaker m1 are shown on the left, m2 on the right.

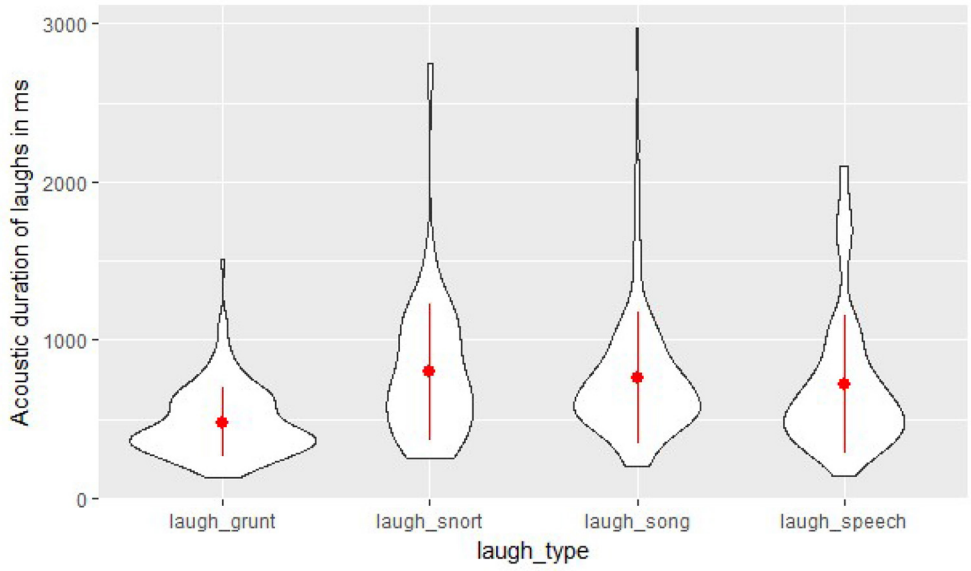


FIGURE 5
Violin plots of acoustic durations (in ms) of the four laughter types. Red dots correspond to the means and red vertical lines to the standard deviations.

Respiratory properties of shared and individual laughter

Similar to the acoustic data, the duration of laughter bouts identified in the respiratory signal was influenced by the laughter type [$F_{(3)} = 5.8$, $p < 0.001$] and the presence of an overlap [$F_{(1)} = 20.4$, $p < 0.001$]. Even though the respiratory data measured longer laugh durations as compared to the acoustic data, the overall patterns were similar across the two measurements. As shown in Figure 6, laughter duration was generally shorter in laughs classified as grunts (~860 ms on average) than in all other laughter categories (~1,100 ms on average; all comparisons with the grunt-like laughs were significant with $t > 2.4$, $p < 0.05$). Moreover, about 340 ms lengthening applied when a laugh was shared (i.e., when the interlocutors' respiratory motions overlapped in time). Unlike the acoustic data, the respiratory

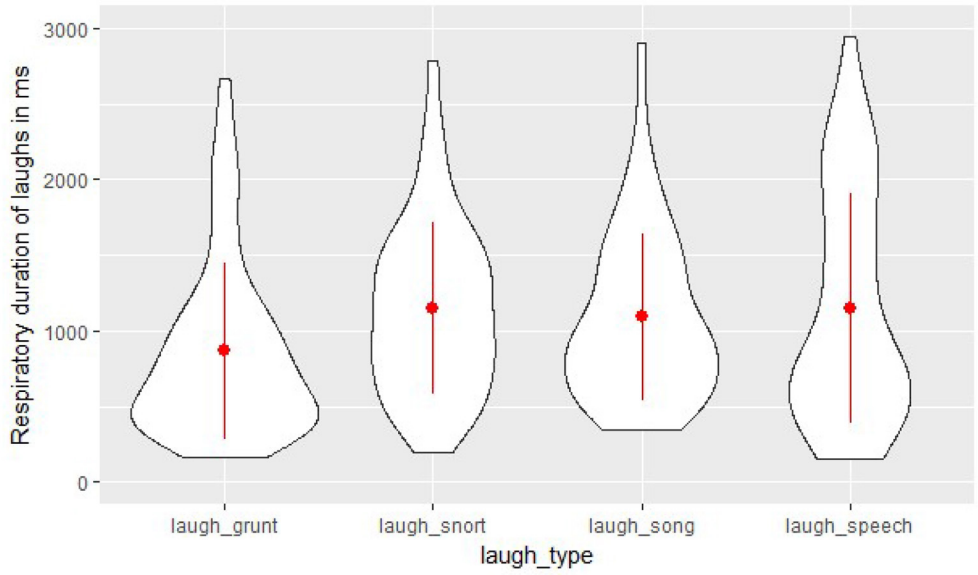


FIGURE 6
Violin plots of respiratory durations (in ms) of the four laughter types under investigation. Red dots correspond to the means and red vertical lines to the standard deviations.

data showed an effect of mutual attraction among daters. Shared laughter defined with reference to the participants' respiration was significantly different across dating pairs. Those pairs who scored higher on mutual attraction also showed longer laughter overlaps [$F_{(1)} = 4.6, p < 0.05$].

Other aspects of respiratory kinematics also varied systematically in response to the factors under investigation. The initial velocity peak depended on the laughter type [$F_{(3)} = 9.8, p < 0.001$]: It was slowest (i.e., closer to zero) in laughs which were categorized as laugh-snort as compared to all other laughter types (all comparisons significant with $t > 2.0, p < 0.05$). The highest velocity peak was observed in laugh-song which was also significantly different from laugh-snort ($t = 5.1, p < 0.001$) and laugh-speech ($t = 2.7, p < 0.01$).

Furthermore, the amplitude of respiratory displacement during laughter was also affected by the laughter type [$F_{(3)} = 6.0, p < 0.001$]. A respiratory displacement was very subtle in laugh-grunts (mean: 7.2 V). In contrast, the displacement was very pronounced in laugh-snorts (mean: 11.4 V) and laugh-songs (mean: 10.2 V), though not all comparisons reached significance at the set alpha level (only the comparison between song-like and grunt-like laughs was significant, $t = 3.4, p < 0.001$).

To check for a fine-tuned overlap of the respiratory signals between the daters, we calculated the time delay between the onsets and the offsets of shared laughter in each pair of speakers (see Figure 3) and tested if these durations were affected by attraction. Against expectation, these analyses did not show any systematic effect of attraction. The finding speaks against a fine-grained view of respiratory synchronization between daters. The predicted effect of an increased mutual attraction on respiration seems to occur on a rather broad level, in terms of a frequent laughter overlap and a longer duration of the overlap. A more fine-grained synchronization account is not supported by these data.

Is there a good predictor for growing attraction?

Following the aims of the study, the above sections addressed the role that attraction may play in laughter behavior during speed dating. In the final section we explored which of the observed behaviors during the date might have contributed to growing mutual attraction during speed dating. An increase in romantic attraction was measured by the subtraction of the score that participants gave their interlocutors prior to the conversation from the score that was given at the end of the 5-min date: The higher the resulting score difference, the higher the potential role of conversational behaviors for the development of attraction. Here, the score varied from −1 (given by m1 to f3) up to 3 (given by f2, f3, f5 to m1). In general, a greater increase in the attraction score after the date was reported by women (1.5 on average) than by men (0.75 on average), though the difference did not reach significance ($V = 37, p = 0.08$, n.s.).

Focusing on the results discussed above, we tested the role of an average laughter duration, percentage of shared laughs, percentage of voiced laughter (i.e., laugh-song vs. all other

laughter types, following Grammer and Eibl-Eibsfeldt's, 1990) and percentage of snort-like laughs (since this was the only laughter type whose occurrence was moderated by the presence of attraction, see Section: Time course and duration of laughter). An ordinal mixed-effects model was fitted to the data and revealed one main effect of an averaged laughter duration ($z = 3.5$, $p < 0.001$). That is, daters who laughed longer were perceived as more attractive at the end of the conversation than at the first sight.

Discussion

The present study was conducted to investigate how non-verbal vocal behavior such as laughter might be a sign of mutual attraction among cisgender, heterosexual interlocutors during speed dating. Different laughter types and their kinematic characteristics were considered. Given the complexity of multi-channel recordings in the present experimental set-up, current results are limited to eight participants and 12 conversations in total and will benefit from a replication with a larger sample and an increased number of male speakers and dyadic conversations. This being said, current findings are in full agreement with previous work (e.g., McFarland, 2001; McFarland et al., 2013; Michalsky and Schoormann, 2018). The observed laughter frequencies are commensurate with the frequencies reported in similar research (e.g., Grammer and Eibl-Eibsfeldt's, 1990; Vettin and Todt, 2004; Kurtz and Algoe, 2015). Moreover, our study adds some new insights to the existing evidence and new perspectives on potential communicative functions of laughter during a romantic communicative interaction.

Laughter and attraction

Based on the present evidence, we may conclude that laughter plays an important role in moderating attraction among mixed-sex, heterosexual dyads, potentially promoting mutual attraction during a speed-date. We combined individual attraction ratings and mutual attraction between the daters in one attraction score. While this score may have some limitations, it might be more robust for statistical analysis than single values. The findings support our hypotheses and reinforce the idea that laughter helps with fostering social and romantic bonds (Kurtz and Algoe, 2015; Bryant et al., 2016; Montoya et al., 2018). Unlike in some previous literature (McAdams et al., 1984; Grammer, 1990), men and women of our study laughed equally frequently, which is likely to be due to the context of their conversation. A dating encounter in our study contrasts with same-sex interviews (McAdams et al., 1984) or non-romantic conversations with a stranger of the opposite sex (Grammer, 1990) in that the main purpose of the short conversations investigated here was to explore a possibility of a romantic bond.

We found many instances of shared laughter in our speed-dating corpus, and their presence was moderated by mutual attraction between the daters. This finding supports our initial hypothesis and is in line with previous research (Kurtz and Algoe, 2015; Bryant et al., 2016; Michalsky and Schoormann, 2018). In the context of the present results, laughter accommodation can be considered only partial (Giles, 1973; Giles et al., 1991; Giles and Ogay, 2007). Specifically, the prediction that shared laughs would be of the same type (cf. Ludusan and Wagner, 2022) was not borne out by the data. Merely 25% of all shared laughs were identified as being of the same category, and their presence was not moderated by attraction.

Instead, our analyses showed that certain laughter types—more specifically, laugh-song (cf. Grammer and Eibl-Eibsfeldt's, 1990) and laugh-snort (as opposed to laugh-speech or laugh-grunt in our classification)—were significantly more likely to invite a joint laughter bout. The two former laughter types differ quite substantially in their temporal, respiratory and intensity features (see Figure 2, cf. Curran et al., 2018). While laugh-song represents a high-sonority, high-energy laughter of the prototypical hahaha-type, laugh-snort with its low energy and relatively low sonority is (at least in terms of its acoustic salience) the exact opposite. The present finding that both of these very different laughs were highly likely to invite a romantic interlocutor to join in is therefore intriguing and somewhat unexpected. Previous research by Curran et al. (2018) investigated what impact laughter intensity might have on the perception of laughter as spontaneous or volitional, showing that high vs. low levels of intensity in laughter cannot be easily mapped onto the dichotomy of spontaneous vs. volitional laughter. Low-intensity laughs being particularly ambiguous with respect to their spontaneity (Wild et al., 2003; Scott et al., 2014; McGettigan et al., 2015). However, research by Lavan et al. (2016) demonstrated that nasality is an extremely salient perceptual feature of deliberate laughs. Such laughs are also rated low on arousal in comparison to a voiced, open-mouthed laugh like laugh-song in our classification (Ruch and Ekman, 2001; Lavan et al., 2016).

It seems that shared laughs observed in our data fall into two categories. On the one hand, laugh-song laughter that is associated with higher arousal and more spontaneity might sound more contagious to the interlocutor, thus more inviting to join in. On the other hand, laugh-snort that has a less aroused and a more volitional quality to it might be experienced as a non-verbal invitation to establishing rapport between the daters, thus playing to the cooperative nature of laughter (cf. Mehu and Dunbar, 2008; Davila-Ross et al., 2011; Curran et al., 2018). These observations warrant a dedicated follow-up study into the communicative functions of different laughter types in the context of dating. Future work on the topic would further benefit from an investigation of the relations between the linguistic content of verbal interactions and laughter occurrence

(Mazzocconi et al., 2020), specifically examining the potential role of humorous content on individual and shared laughs (cf. Bachorowski et al., 2001; Holt, 2019).

In our data, men frequently initiated a shared laugh, while the likelihood of female laughter sparking a laugh from a male interlocutor was much lower. These results are very comparable with the gender effects reported for shared laughs in long-term couples by Kurtz and Algoe (2015) and are in line with the previous research on humor and attraction (Bressler et al., 2006). This research has shown that women tend to favor men who make them laugh, while men are more attracted to women who laugh at their jokes. Our study provides unique evidence on how these gender preferences may play out in the context of an emerging attraction and suggest that laughter fulfills similar functions at the beginning of a potential relationship, as it does in established couples (Kurtz and Algoe, 2015).

Our study further showed that shared laughter prolonged the overall duration of a laugh and that longer laughs of an interlocutor increased the sense of an attraction to them. Previous research showed that duration of a laugh plays a role for its perceptual interpretation, with a longer laugh being judged as more spontaneous and authentic (Bryant and Aktipis, 2014; Lavan et al., 2016). In contrast, none of the factors in the present investigation influenced the timings of the beginning and the end of shared laughter, though latencies in both onsets and offsets of joint laughs fell into a comparable window of 320–360 ms. These latencies align with the delays observed in shadowing (Marslen-Wilson, 1973) and might be indicative of the time course involved in some general mental processes that are not specific to laughter (cf. Marslen-Wilson, 1985). Discourse analysts suggest that interactional contributions within dialogues follow strict timing patterns (Kendrick and Torreira, 2015), and for example delays from an accepted response window may be interpreted as an indication of a negative response to a posed question (e.g., Clayman, 2002). Future work on timing patterns of laughter will help to uncover if the alignment between an initial and a burst of incoming laughter can increase, or decrease, the sense of attraction and rapport between the interlocutors.

Respiratory features of laughter

To date, only a limited number of studies have addressed respiratory kinematics of laughter in social interactions. The core advantage of this method is that physiology of speech offers a rich source of information about the speaker but does not necessarily leave an acoustic imprint in the speech signal (e.g., Lisker, 1986; Pouplier and Goldstein, 2005; Schaeffler et al., 2015; Cleland et al., 2019). Compared to the acoustic data discussed in Section: Distribution of shared and individual laughter, our respiration findings were consistent with the overall patterns of laughter and its role in attraction. For example, shared laughs were longer in acoustics and respiration. The subtle effect of attraction on shared laughter duration could not be ascertained in acoustics, while the respiratory laughter overlap was clearly longer in pairs who scored higher on attraction. This signifies the importance of physiological data for understanding social interaction. Overall, the physiologically defined duration of laughter exhalation was longer than the acoustically measured durations. Such differences may arise from the fact that periods of acoustic silence during laughter preparation or at the end of a laughter bout can correspond to some ongoing respiratory activity identifiable in the physiological signal only. Effects like these have been well-documented for many acoustically silent periods during speech production when the articulators are moving but not producing any audible acoustic cues (e.g., closure intervals in stop production, cf. Lisker, 1986, or during pauses, e.g., Ramanarayanan et al., 2009; Krivokapić et al., 2020).

Previous work by McFarland (2001) focused on respiratory kinematics in different conversational modes (speaking, listening, turn-taking) in same-sex dyads. All participants were familiar with each other and had been close friends for at least 3 years prior to the study. That is, romantic attraction was an extremely unlikely factor in those dialogues. McFarland et al. (2013) results based on cross-correlation showed a high degree of in-phase-coordination in respiratory kinematics between conversational partners during their laughter, especially in the proximity of a speaker change (i.e., ±5 s before or after turn taking). In contrast, participants of the 12 conversational interactions studied here met each other for the first time and participated in a speed dating experiment where attraction, or the lack thereof, was expected to have an influence on the duration of respiratory overlap while laughing together. The kinematic measures employed here were more fine-grained than those obtained in McFarland et al. (2013) study, though we also examined only the onset and offset of the breathing cycles in shared laughter. Using these measures, we did not observe an effect of attraction on the temporal synchronization at the beginning or the end of joint laughter. However, further analyses of kinematic signals ought to examine possible effects of attraction on the magnitude of respiratory movements during shared laughter.

The present study focused primarily on the exhalation phase as it has been recognized as one of the most important phases in laughter (Filippelli et al., 2001). During exhalation in laughter, the lung volume tends to decrease rapidly, and according to the visual inspections of these data, the volume sometimes reached a much lower level than is typical of speech production. Such rapid respiratory changes are likely to depend on the type of laughter—a variable which was not assessed in previous research. In line with our hypothesis, the data reveal an effect of laughter type on the properties of the exhalation. More specifically, laugh-grunts were produced with the shortest duration and a smallest displacement as compared to all other laugh types. The two

respiratory parameters are clearly linked. That is, at any given speed of respiratory motion, shorter duration sets limits to the degree of a respiratory displacement and will result in smaller displacements. However, the speed of respiratory motion, as measured by the initial velocity peak in the exhalation phase, was not particularly slow in laugh-grunts. On the contrary, the slowest peak was measured in laugh-snorts, while the fastest peak occurred in laugh-songs. These findings contradict our initial, physiologically motivated account of laughter, which predicted high velocity peaks to occur in voiceless laughs when the glottis is open, and the airway is not obstructed by the vocal folds. If this hypothesis were to have found support, we would have observed the opposite patterns. That is, laugh-snorts and laugh-grunts would have measured higher velocity peaks, while laugh-songs would have been produced with the lowest peaks. However, obstructions of air flow may not only occur at the glottal level, but also elsewhere in the vocal tract. Against the background of the current results, future research on the respiratory physiology of laughter should consider laughter intensity and the corresponding subglottal pressure, along with phonation and aerodynamics.

Summary, conclusion, and outlook

Our small-scale study combined acoustic, physiological and psychological measurements of opposite-sex interlocutors involved in romantic conversations during speed dating. Partially supporting the ideas of the Accommodation Theory that non-verbal conversation behaviors can show convergence (Giles and Ogay, 2007), we found that high-attraction dates had a higher number of shared laughs with longer kinematic overlaps, though these laughs were not of the same type, as had been predicted. Laugh-snorts were more prevalent in high-attraction dates, while laugh-songs were more likely to induce a joint laughter bout, leading to the conclusion that a nuanced view of laughter types would benefit the understanding of the many functions that laughter might serve in (romantic) communicative interactions. Importantly, we would like to highlight that our study, along with much previous research, is limited to cisgender, heterosexual behaviors during speed dating. The present findings may not generalize to other contexts of mutual attraction in conversational dyads. More work is needed to shed light on the social importance of laughter across diverse dating contexts.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. The patients/participants provided their written informed consent to participate in this study.

Author contributions

TR and SF designed the experiment, ran the recording, jointly annotated the data, and wrote and revised the manuscript in response to the reviewer comments. TR analyzed the acoustic data and conducted the statistical analyses. SF analyzed the respiratory data. Both authors contributed to the article and approved the submitted version.

Funding

This work was supported by an intramural travel grant from the University of Kent to TR and financial support of the Leibniz Association to SF.

Acknowledgments

We would like to thank our participants who were amazingly open, authentic, friendly and extremely supportive of the nervous experimenter crew throughout the duration of the experiment. We would also like to thank Jörg Dreyer for the technical assistance, Olivia Maky for the support during data collection and preparation, and James Brand for a good laugh. We are further grateful to Dani Byrd, one reviewer and the issue editor for a helpful and constructive discussion during the publication process. An earlier version of this study along with preliminary analyses of the data was discussed at the laughter workshop 2018 in Paris, which also published short conference proceedings.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated

organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Bachorowski, J.-A., Smoski, M. J., and Owren, M. J. (2001). The acoustic features of human laughter. *J. Acoust. Soc. Am.* 110, 1581–1597. doi: 10.1121/1.1391244

Bates, D., Maechler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* 67, 1–48. doi: 10.18637/jss.v067.i01

Bloch, S., Lemeignan, M., and Aguilera, N. (1991). Specific respiratory patterns distinguish among human basic emotions. *Int. J. Psychophysiol.* 11, 141–154.

Boersma, P., and Weenink, D. (2018). *Praat: Doing Phonetics by Computer [Computer Program]. Version 6.0.37*. Available online at: http://www.praat.org/ (accessed February, 2018).

Bressler, E. B., Martin, R. A., and Balshine, S. (2006). Production and appreciation of humour as sexually selected traits. *Evol. Hum. Behav.* 27, 121–130. doi: 10.1016/j.evolhumbehav.2005.09.001

Bryant, G. A., and Aktipis, C. A. (2014). The animal nature of spontaneous human laughter. *Evol. Hum. Bsehav.* 35, 327–335. doi: 10.1016/j.evolhumbehav.2014.03.003

Bryant, G. A., Fessler, D. M. T., Fusaroli, R., Clint, E., Aarøe, L., Apicella, C. L., et al. (2016). Detecting affiliation in colaughter across 24 societies. *Proc. Natl. Acad. Sci. U. S. A.* 113, 4682–4687. doi: 10.1073/pnas.1524993113

Christensen, R. H. B. (2018). *Package 'Ordinal'*. Available online at: https://cran.r-project.org. (accessed May 08, 2021).

Clayman, S. E. (2002). "Sequence and solidarity," in *Group Cohesion, Trust and Solidarity*, eds S. R. Thye, and E. J. Lawler (Oxford: Elsevier Science Ltd.), 229–253. doi: 10.1016/S0882-6145(02)19009-6

Cleland, J., Scobbie, J. M., Roxburgh, Z., Heyde, C., and Wrench, A. (2019). Enabling new articulatory gestures in children with persistent speech sound disorders using ultrasound visual biofeedback. *J. Speech Lang. Hear. Res.* 62, 1–18. doi: 10.1044/2018_JSLHR-S-17-0360

Curran, W., McKeown, G. J., Rychlowska, M., Elisabeth André, E., Wagner, J., and Lingenfelser, F. (2018). Social context disambiguates the interpretation of laughter. *Front. Psychol.* 8, 2342. doi: 10.3389/fpsyg.2017.02342

Davila-Ross, M., Allcock, B., Thomas, C., and Bard, K. A. (2011). Aping expressions? Chimpanzees produce distinct laugh types when responding to laughter of others. *Emotion* 11, 1013–1020. doi: 10.1037/a0022594

Feleky (1915). *The influence of the emotions on respiration* (M.A. thesis). Columbia University, New York, NY, United States.

Filippelli, M., Pellegrino, R., Iandelli, I., Misuri, G., Rodarte, J. R., Duranti, R., et al. (2001). Respiratory dynamics during laughter. *J. Appl. Physiol.* 90, 1441–1446. doi: 10.1152/jappl.2001.90.4.1441

Finkel, E. J., and Eastwick, P. W. (2008). Speed dating. *Curr. Direct. Psychol. Sci.* 17, 193–197. doi: 10.1111/j.1467-8721.2008.00573.x

Finkel, E. J., Eastwick, P. W., and Matthews, J. (2007). Speed-dating as an invaluable tool for studying romantic attraction: a methodological primer. *Pers. Relationsh.* 14, 149–166. doi: 10.1111/j.1475-6811.2006.00146.x

Giles, H. (1973). Accent mobility: a model and some data. *Anthropol. Linguist.* 15, 87–109.

Giles, H., Coupland, J., and Coupland, N. (1991). *Contexts of Accommodation: Developments in Applied Sociolinguistics*. Cambridge: CUP.

Giles, H., and Ogay, T. (2007). "Communication accommodation theory," in *Explaining Communication: Contemporary Theories and Exemplars*, eds B. B. Whaley, and W. Samter (Mahwah, NJ: Laurence Erlbaum), 293–310.

Glenn, P. (1989). Initiating shared laughter in multi-party conversations. *West. J. Speech Commun.* 53, 127–149.

Glenn, P. (1992). Current speaker initiation of two-party shared laughter. *Res. Lang. Soc. Interact.* 25, 139–162.

Grammer, K. (1990). Strangers meet: laughter and nonverbal signs of interest in opposite-sex encounters. *J. Nonverb. Behav.* 14, 209–236.

Grammer, K., and Eibl-Eibsfeldt, I. (1990). "The ritualisation of laughter," in *Die Natürlichkeit der Sprache und der Kultur*, ed W. A. Koch (Bochum: Brockmeyer), 192–214.

Hay, J., and Drager, K. (2010). Stuffed toys and speech perception. *Linguistics* 48, 865–892. doi: 10.1515/ling.2010.027

Hay, J., Podlubny, R., Drager, K., and Mcauliffe, M. (2017). Car-talk: location-specific speech production and perception. *J. Phonet.* 65, 94–109. doi: 10.1016/j.wocn.2017.06.005

Hixon, T. J., Goldman, M. D., and Mead, J. (1973). Kinematics of the chest wall during speech production: volume displacements of the rib cage, abdomen, and lung. *J. Speech Hear. Res.* 16, 78–115.

Holt, E. (2019). Conversation analysis and laughter. *Concise Encycl. Appl. Linguist.* 275, 275–279. doi: 10.1002/9781405198431.wbeal0207.pub2

Jaffe, J., and Feldstein, S. (1970). *Rhythms of Dialogue*. New York, NY: Academic Press.

Jefferson, G. (1979). "A technique for inviting laughter and its subsequent acceptance/declination," in *Everyday Language: Studies in Ethnomethodology*, ed G. Psathas (New York, NY: Irvington Publishers), 79–96.

Jefferson, G. (2004). A note on laughter in 'male-female' interaction. *Discourse Stud.* 6, 117–133. doi: 10.1177/1461445604039445

Katsika, A., Krivokapić, J., Mooshammer, C., Tiede, M., and Goldstein, L. (2014). The coordination of boundary tones and its interaction with prominence. *J. Phonet.* 44, 62–82. doi: 10.1016/j.wocn.2014.03.003

Kendrick, K. H., and Torreira, F. (2015). The timing and construction of preference: a quantitative study. *Discourse Processes* 52, 255–289. doi: 10.1080/0163853X.2014.955997

Kipper, S., and Todt, D. (2001). Variation of sound parameters affects the evaluation of human laughter. *Behaviour* 138, 1161–1178. doi: 10.1163/156853901753287181

Kohler, K. J. (2008). 'Speech-smile', 'speech-laugh', 'laughter' and their sequencing in dialogic interaction. *Phonetica* 65, 1–18. doi: 10.1159/000130013

Krivokapić, J., Styler, W., and Parrell, B. (2020). Pause postures: the relationship between articulation and cognitive processes during pauses. *J. Phonet.* 79, 100953. doi: 10.1016/j.wocn.2019.100953

Kurtz, L., and Algoe, S. (2015). Putting laughter in context: shared laughter as behavioral indicator of relationship well-being. *J. Int. Assoc. Relationsh. Res.* 22, 573–590. doi: 10.1111/pere.12095

Kuznetsova, A., Brockhoff, P. B., and Christensen, R. H. B. (2017). lmerTest package: tests in linear mixed effects models. *J. Stat. Softw.* 82, 1–26. doi: 10.18637/jss.v082.i13

Lavan, N., Scott, S., and McGettigan, C. (2016). Laugh like you mean it: authenticity modulates acoustic, physiological and perceptual properties of laughter. *J. Non-verb. Behav.* 40, 133–149. doi: 10.1007/s10919-015-0222-8

Lefcourt, R. M. (2000). *Humor: The Psychology of Living Buoyantly*. New York, NY: Plenum Publishers.

Lisker, L. (1986). "Voicing" in english: a catalogue of acoustic features signalling /b/ versus /p/ in trochees. *Lang. Speech* 29, 3–11.

Ludusan, B., and Wagner, P. (2022). Laughter entrainment in dyadic interactions: temporal distribution and form. *Speech Commun.* 136, 42–52. doi: 10.1016/j.specom.2021.11.001

Marslen-Wilson, W. (1973). Linguistic structure and speech shadowing at very short latencies. *Nature* 244, 522–523.

Marslen-Wilson, W. (1985). Speech shadowing and speech comprehension. *Speech Commun.* 4, 55–73.

Mazzocconi, C., Tian, Y., and Ginzburg, J. (2020). What's your laughter doing there? A taxonomy of the pragmatic functions of laughter. *IEEE Trans. Affect. Comput.* 1–19. doi: 10.1109/TAFFC.2020.2994533

McAdams, D. P., Jackson, R. J., and Kirshnit, C. (1984). Looking, laughing, and smiling in dyads as a function of intimacy motivation and reciprocity. *J. Personal.* 52, 261–273.

McFarland, D. H. (2001). Respiratory markers of conversational interaction. *J. Speech Lang. Hear. Res.* 44, 128–143. doi: 10.1044/1092-4388 (2001/012)

McFarland, D. H., Jurafsky, D., and Rawlings, C. (2013). Making the connection: social bonding in courtship situations. *Am. J. Sociol.* 118, 1596–1649. doi: 10.1086/670240

McGettigan, C., Walsh, E., Jessop, R., Agnew, Z. K., Sauter, D. A., Warren, J. E., et al. (2015). Individual differences in laughter perception reveal roles for mentalizing and sensorimotor systems in the evaluation of emotional authenticity. *Cerebral Cortex* 25, 246–257. doi: 10.1093/cercor/bht227

Mehu, M., and Dunbar, R. I. M. (2008). Naturalistic observations of smiling and laughter in human group interactions. *Behaviour* 145, 1747–1780. doi: 10.1163/156853908786279619

Michalsky, J., and Schoormann, H. (2018). "Phonetic entrainment of laughter in dating conversations: on the effects of perceived attractiveness and conversational quality," in *Proceedings of Laughter Workshop*, eds J. Ginzburg, and C. Pelachaud (Paris: Sorbonne Université Paris), 40–44.

Michalsky, J., Schoormann, H., and Niebuhr, O. (2017). Turn transitions as salient places for social signals–local prosodic entrainment as a cue to perceived attractiveness and likability," in *Proceedings of the Conference on Phonetics and Phonology of German-Speaking Areas* (Berlin), 125–128.

Montoya, R. M., Kershaw, C., and Prosser, J. L. (2018). A meta-analytic investigation of the relation between interpersonal attraction and enacted behaviour. *Psychol. Bull.* 144, 673–709. doi: 10.1037/bul0000148

Nwokah, E. E., Hsu, H.-C., Davies, P., and Fogel, A. (1999). The integration of laughter and speech in vocal communication: a dynamic systems perspective. *J. Speech Lang. Hear. Res.* 42, 880–894.

Owren, M., and Bachorowski, J. (2003). Reconsidering the evolution of nonlinguistic communication: the case of laughter. *J. Nonverb. Behav.* 27, 183–200. doi: 10.1023/A:1025394015198

Pouplier, M., and Goldstein, L. (2005). Asymmetries in the perception of speech production errors. *J. Phonet.* 33, 47–75. doi: 10.1016/j.wocn.2004.04.001

Provine, R. R. (1996). Laughter. *Am. Sci.* 84, 38–45.

Provine, R. R. (2000). *Laughter: A Scientific Investigation*. New York, NY: Viking.

Provine, R. R. (2004). Laughing, tickling, and the evolution of speech and self. *Curr. Direct. Psychol. Sci.* 13, 215–218. doi: 10.1111/j.0963-7214.2004.00311.x

Provine, R. R., and Emmorey, K. (2006). Laughter among deaf signers. *J. Deaf Stud. Deaf Educ.* 11, 403–409. doi: 10.1093/deafed/enl008

Ramanarayanan, V., Bresch, E., Byrd, D., Goldstein, L., and Narayanan, S. S. (2009). Analysis of pausing behavior in spontaneous speech using real-time magnetic resonance imaging of articulation. *J. Acoust. Soc. Am.* 126, EL160–EL165. doi: 10.1121/1.3213452

Ruch, W., and Ekman, P. (2001). "The expressive pattern of laughter," in *Emotion, Qualia and Consciousness*, ed A. Kaszniak (Tokyo: World Scientific), 426–443. doi: 10.1142/9789812810687_0033

Sauter, D. A., Eisner, F., Ekman, P., and Scott, S. K. (2010). Cross-cultural recognition of basic emotions through nonverbal emotional vocalizations. *Proc. Natl. Acad. Sci. U. S. A.* 107, 2408–2412. doi: 10.1073/pnas.0908239106

Schaeffler, S., Scobbie, J. M., and Schaeffler, F. (2015). "Complex patterns in silent speech preparation," in *Proceedings of the 18th International Congress of Phonetic Sciences*, ed The Scottish Consortium for ICPhS 2015 (Glasgow: The University of Glasgow). ISBN 978-0-85261-941-4. Paper number 866, 1–5. Available online at: https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2015/Papers/ICPHS0866.pdf (accessed August 05, 2021).

Scott, S. K., Lavan, N., Chen, S., and McGettigan, C. (2014). The social life of laughter. *Trends Cognit. Sci.* 18, 618–620. doi: 10.1016/j.tics.2014.09.002

Smoski, M. J., and Bachorowski, J.-A. (2003). Antiphonal laughter between friends and strangers. *Cognit. Emot.* 17, 327–340. doi: 10.1080/02699930302296

Tanaka, H., and Campbell, N. (2011). "Acoustic features of four types of laughter in natural conversational speech," in *Proceedings of the 17th International Congress of Phonetic Sciences* (Hong Kong), 1958–1961.

Trouvain, J. (2003). "Segmenting phonetic units in laughter," in *Proceedings of the 15th International Congress of Phonetic Sciences Barcelona* (Adelaide: Causal Productions), 2793–2796.

Truong, K. P., and Trouvain, J. (2014). "Investigating prosodic relations between initiating and responding laughs," in *Proceedings of the 15th Annual Conference of the International Speech Communication Association (Interspeech)*, 1811–1815. doi: 10.21437/Interspeech.2014-412

Vettin, J., and Todt, D. (2004). Laughter in conversation: features of occurrence and acoustic structure. *J. Nonverb. Behav.* 28, 93–115. doi: 10.1023/B:JONB.0000023654.73558.72

Wild, B., Rodden, F. A., Grodd, W., and Ruch, W. (2003). Neural correlates of laughter and humour. *Brain* 126, 2121–2138. doi: 10.1093/brain/awg226

Appendix

Table A1 Initial (pre) and final (post) scores given by the two male speakers (in rows) to the 6 female speakers (in columns).

Dater	f1 pre	f1 post	f2 pre	f2 post	f3 pre	f3 post	f4 pre	f4 post	f5 pre	f5 post	f6 pre	f6 post
m1	6	6	7	8	7	6	7	9	7	9	7	7
m2	4	4	8	9	5	6	6	7	8	9	8	9

Table A2 Initial (pre) and final (post) scores given by the six female speakers (in rows) to the two male speakers (in columns).

Dater	m1 pre	m1 post	m2 pre	m2 post
f1	6	7	2	3
f2	4	7	2	3
f3	2	5	2	2
f4	5	7	3	4
f5	3	6	3	5
f6	5	6	2	2

Table A3 Scores for mutual interest in a second date: 2 = both like to meet again, 1: one interlocutor likes to meet again, but not the other, 0: none is interested in a further meeting.

Dating pairs	Further interest in a date
m1-f1	1
m1-f2	2
m1-f3	1
m1-f4	1
m1-f5	1
m1-f6	1
m2-f1	0
m2-f2	1
m2-f3	0
m2-f4	1
m2-f5	1
m2-f6	1

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 05 August 2022
DOI 10.3389/fpsyg.2022.966254

Check for updates

OPEN ACCESS

EDITED BY
Oliver Niebuhr,
University of Southern Denmark, Denmark

REVIEWED BY
Steven Bellman,
University of South Australia, Australia
Felix Septianto,
The University of Queensland, Australia
Carla Canestrari,
University of Macerata, Italy

*CORRESPONDENCE
Dan Li
Lidan@stu.zuel.edu.cn

SPECIALTY SECTION
This article was submitted to Emotion Science, a section of the journal Frontiers in Psychology

RECEIVED 10 June 2022
ACCEPTED 21 July 2022
PUBLISHED 05 August 2022

CITATION
Ning Y-m, Hu C, Tu T-t and Li D (2022) Offensive or amusing? The study on the influence of brand-to-brand teasing on consumer engagement behavioral intention based on social media.
Front. Psychol. 13:966254.
doi: 10.3389/fpsyg.2022.966254

COPYRIGHT
© 2022 Ning, Hu, Tu and Li. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Offensive or amusing? The study on the influence of brand-to-brand teasing on consumer engagement behavioral intention based on social media

Yu-mei Ning[1,2], Chuan Hu[2], Ting-ting Tu[2] and Dan Li[2]*

[1]Business School, Yulin Normal University, Yulin, China, [2]School of Business Administration, Zhongnan University of Economics and Law, Wuhan, China

With the development of social media, advertising has migrated from traditional media to social media. Marketers are increasingly using social media's brand pages to actively create humorous dialogue interactions with other brands for brand communication to achieve positive business outcomes. Especially brand-to-brand's aggressive humor dialogue can also be an effective brand communication strategy. Based on benign violation theory, we have studied the influence mechanism and boundary condition of the brand-to-brand's aggressive humor styles (low-aggressive and high-aggressive) on consumer engagement behavioral intention in social media context. Through experiments, it is indicated that low-aggressive humor could promote consumer engagement behavioral intention more than high-aggressive humor. Benign appraisal mediates the relationship between low-aggressive humor and consumer engagement behavioral intention. Furthermore, brand personality not only moderates the effect of low-aggressive humor on consumer engagement behavioral intention, but also moderates the mediating role of benign appraisal between low-aggressive humor and consumer engagement behavior intention.

KEYWORDS

aggressive humor, brand dialogue, consumer engagement behavior, brand personality, social media

Introduction

With the popularity of the Internet, the number of users of social media continues to rise, and social media has taken up a large part of people's daily life. Since social media can satisfy most people's entertainment needs, "humor" has gradually become a prominent element of online communication on social media. Social media users are

engaging with humorous content more than ever by spreading the humorous content to others through the network (Shifman, 2012). In addition to interpersonal interactions, brands also use playful entertainment content in anthropomorphic ways to promote consumer-brand interactions (Nielsen, 2015; Cheung et al., 2019). Among them, Durex and Burger King are typical examples of the successful use of anthropomorphic expressions to carry out humorous dialogues with other brands on social media. Durex@Wrigley Gum: "Honey, thank you for being on my left for so many years and being an excuse to buy me." Burger King @ McDonald's: "Every king needs a clown, happy 50th birthday!" Are brands using their brand pages on social media to actively create humorous dialogues and interactions with other brands in an ad that is attempting to engage consumers in the ad's interactions, or is it just to entertain them? How effective are this brand's humorous dialogues and interactions with other brands in motivating consumers to respond to what the brand is promoting in advertising?

Existing research on humorous communication has mainly explored the impact of humorous advertising (the humorous interaction between brands and consumers) on consumers. For example, many scholars have studied traditional media such as television and print media (Eisend, 2009; Warren et al., 2019) and online media (Dessart et al., 2015; Nair, 2016) humorous advertisements have a positive impact on consumers by breaking through the clutter of advertisements. The same is a humorous advertisement, and its humorous style will also be different, some funny and affiliative, and some witty and aggressive. The aggressive humorous posts and memes on social media are one of the most widely spread by users (Taecharungroj and Nueangjamnong, 2014). Research on aggressive humor advertising has pointed out that disparaging humorous television ads increase brand attitudes and advertising recall by increasing the superiority of those who seek high power (Newton et al., 2016). Gregory et al. (2019) argued that aggressive humorous television ads affect advertising effectiveness through individualistic culture's perception of humor. However, only a few studies have organically combined aggressive humor and brand-to-brand interaction to explore the impact of brands' aggressive humor dialogue interaction with other brands on consumers. Exploring this gap is important, as compared with traditional competing advertising, brand-to-brand dialogue is more likely to improve consumers' perception of freedom and be more easily accepted by them (Thomas and Fowler, 2021). Meanwhile, aggressive humorous exchanges are desired by social media users and are commonly considered as the goal of some brand-to-brand interactions (Jargon, 2017), we combine aggressive humor and brand-to-brand interaction. Thomas and Fowler (2021) argued that witty brand teasing can be an effective brand communication strategy. Brand-to-brand low-aggressive humor (relative to high-aggressive humor) increases consumers' interest in the brand initiating the dialogue by reducing consumers' perceptions of manipulative intent. And the response to the dialogue brand needs to select the appropriate humor type according to the type of humor that initiated the dialogue brand, so that the communication effect will be better. This study explored the impact of low-aggressive humor from a negative cognitive perspective (i.e., perceptions of manipulative intent), ignoring the analysis of positive cognitive factors. Because this kind of aggressive competitive advertising will not only generate negative reviews from consumers, but also positive reviews. Even competing ads will have more positive reviews (e.g., perceived brand distinctiveness) (Berendt et al., 2018).

As the dominant theory of explaining the mechanism of aggressive humor's effect, benign violation theory holds that humor occurs when a stimulus is evaluated as containing violation and also the violation is evaluated as benign. Individuals' perception of violation and the degree of benignity will affect their perception of humor and their attitude toward the initiator of humor (Warren and McGraw, 2015). In this theoretical framework, aggressive humor may be considered a harmful offense, or it may trigger laughter, and whether the attack can be evaluated as benign directly determines whether the attack can trigger humor. The achievement of benign attack appraisal is affected by humor initiator and audience factors (Li and Wang, 2022). Meanwhile, consumers' evaluation of brand-initiated humorous communication largely depends on the initiator (Romero and Cruthirds, 2006). Therefore, this research used the benign violation theory, based on social media context, with positive cognitive factors (i.e., benign appraisal) as the mediating variable, and the characteristics of the brand initiating the dialogue (i.e., brand personality) as the moderating variable, to study the influence of the brand's aggressive humor style to other brands on consumer engagement behavioral intention, so as to provide implications for brand's use of effective humor style on social media to bring brand communication and consumer engagement behavior.

Theoretical review and basic hypothesis

Humor and interpersonal humor style

Broadly speaking, humor includes any interesting communication that generates positive emotions and cognitions in individuals, groups, or organizations (Romero and Cruthirds, 2006). Humor has received increasing attention in organizational environment (Romero and Cruthirds, 2006) and advertising strategies (Eisend, 2009).

Importantly, humor is a double-edged sword. For this reason, Martin et al. (2003), for different forms of CEO humor, based on other-directed humor, divided into two styles: positive humor (affiliative humor) and negative humor

(aggressive humor). (1) Affiliative humor refers to a form of benign humor that rewards others and aims to support others to enhance relationships with others. Praise is an example of affiliative humor; (2) aggressive humor refers to harmful humor forms that are entertained by belittling others. Teasing, whining, ridicule, and sarcasm are examples of aggressive humor (Martin et al., 2003). According to the level of aggression, aggressive humor can be divided into low-aggressive humor and high-aggressive humor (Wright and Roloff, 2013). For low-aggressive humor, the attack or disparagement component is less severe (or more lighthearted) (Speck, 1991), and in many cases the tension generated by disparaging another is a mixture of enjoyment tinged with anxiety or guilt over enjoyment of the disparagement (Beard, 2008). Thus, low-aggressive humor is described as prosocial teasing, including flirtatious teasing (Shapiro et al., 1991) and jocular mockery (Haugh, 2014). Conversely, high-aggressive humor relies on disparaging another and is a form of negative comedy designed to improve personal well-being by disparaging another person or group (Warren et al., 2018). High-aggressive humor meant dark, more acerbic in tone, and meant to wound rather than just amuse (Holbert et al., 2011). Thus, high-aggressive humor is described as antisocial teasing, including sarcasm (Ducharme, 1994), ridicule (Billig, 2005), and bullying (Smith et al., 2008).

Brand-to-brand aggressive humor style

Since brands are like other users on social media, dialogue interactions can be conducted in front of actual or potential audiences in public. Therefore, through the anthropomorphic expression of the brand, consumers will regard the brand as a person, and use the cognitive method of interpersonal communication to form a perception of the brand (Kwak et al., 2015). Therefore, the concept of interpersonal humor proposed by existing research can be used for brand dialogue. For the purposes of the study, this research defined brand-to-brand low-aggressive humor (e.g., teasing) as: brand-generated content involving harmful humor directed at an identifiable target competing brand, having fun in a way that disparages the identifiable target competing brand, and the attack or disparagement component are less severe (or more lighthearted). Brand-to-brand high-aggressive humor (e.g., ridicule) is defined as: brand-generated content involving harmful humor directed at an identifiable target competing brand, having fun in a way that disparages the identifiable target competing brand, and the attack or disparagement component are severe (or more extreme). Since the inter-brand dialogue involves the subject brand of the dialogue interaction (the brand that initiates the dialogue) and the object brand of the dialogue interaction (the identifiable target competing brand), but compared with the object brand of the dialogue interaction, the subject brand of the dialogue interaction plays the leading role in the inter-brand dialogue, so that their use of humorous dialogue strategies has a greater effect on consumers (Romero and Cruthirds, 2006). Therefore, this study only focuses on the research of the subject brand of the dialogue interaction, that is, to explore how the humor style of the subject brand of the dialogue interaction affects consumers' perception and behavioral intention of the subject brand of the dialogue interaction, but does not include research on consumers' evaluation of the object brand of dialogue interaction.

Aggressive humor type and consumer engagement behavioral intention

Consumer engagement in social media context encompasses affective, cognitive, and behavioral dimensions. Importantly, user engagement is defined as behavioral manifestations toward a brand which are expressions of underlying psychological states resulting from a consumer's interactive relationship with a brand (Brodie et al., 2011). Consumer engagement is usually measured through likes, comments, and shares (Coelho et al., 2016). Likes are a way for consumers to react to brand-initiated posts by not only showing consumers endorsement of brand-initiated posts, but also giving them credibility. Comments are a way for consumers to provide their comments on posts by adding new content. Reviews can help companies maintain the conversation, but also give consumers the power to organize the conversation. Retweets are a channel for consumers to spread brand information. Consumers' sharing expands the reach of brand information's audience and enhances interactivity.

On social media, novel, smart and delightful content generates higher audience responses (Brennan et al., 2015). Funny, playful ads help develop relationships by promoting conversation with others, which in turn are better for consumers to share (Campo et al., 2013). Due to the aggressive characteristics of aggressive humor, it will evoke high arousal emotions (e.g., awe, surprise, anger, and anxiety) in the audience (Taecharungroj and Nueangjamnong, 2014), and many scholars found a strong relationship between high arousal emotions and virality (Kaplan and Haenlein, 2011; Berger and Milkman, 2012). Brand teasing or ridicule creates a sense of connection with audiences. Because audiences feel that the brand teasing or ridicule was created for their benefit, they respond by expressing appreciation. This appreciation can take the form of laughter or even involvement in the process of continuing to tease or ridicule.

In the social media context, brands have fun by disparaging identifiable target competing brands. When disparagement reaches extremes, as is the case of high-aggressive humor, it may be viewed as excessive to the point of hostility (Martin et al., 2003). Further, humorous advertisements that trigger negative emotional responses can lead to public complaints

(Beard, 2008). Consumers may prefer other less offensive forms of humor (Speck, 1991). Therefore, high-aggressive humor may attenuate the positive effects of humor. In addition, Holbert et al. (2011) showed the potential divergent effects of low-aggressive and high-aggressive humor on persuasion in the context of political campaign. Thomas and Fowler (2021) argued that when brands use low-aggressive humor to dialogue with another brands, consumers show higher interest in the brand that initiated the dialogue. Thus, compared with high-aggressive humor, low-aggressive humor is more beneficial to increase consumers' interest in the brand, and consumers' intention to comment and share.

Therefore, we have put forward Hypothesis H1: Compared with high-aggressive humor, low-aggressive humor generates more consumer engagement behavioral intention.

The mediating effect of benign appraisal

The benign violation theory is a theory that has been used in recent years to explain how humor occurs and affects its surroundings. The theory proposes that humor occurs when something that is perceived to threaten a person's well-being, identity, or normative belief structure simultaneously seems okay, acceptable, non-threatening, harmless or inconsequential (i.e., benign appraisal) (McGraw and Warren, 2010; McGraw et al., 2012b). Individuals' perceptions of violations and the degree of benignity will affect their perception of humor and their attitudes toward the humor initiator (Warren and McGraw, 2015). In other words, the perception that something that is wrong is actually okay can transform an otherwise negative experience to a positive experience characterized by laughter and amusement (Apter, 1982). Excessive violations of the norm, however, offend or threaten the perceiver, thereby inhibiting the effect of humor (McGraw and Warren, 2010). Norms are defined as perceptions, attitudes, and behaviors that a group or society approves of and abides by Baron et al. (1992). Violations include not only threats to physical well-being (e.g., tickling) but also identity threats (e.g., teasing, sarcasm) and behaviors that violate cultural norms (e.g., inappropriate dress), social norms (e.g., strange behavior), moral norms (e.g., disrespectful behavior), communication norms (e.g., puns, malapropisms) and logic norms (e.g., absurdities, non-sequiturs). Benign appraisal refers to an individual's subjective perception of normative, acceptable, and sensible things, that is, a benign violation perception (McGraw and Warren, 2010).

Thus, according to the benign violation theory, brands disparaging identifiable target competing brands can be viewed by consumers as a violation of expected existing communication norms. Whether the disparagement can be evaluated as benign directly determines whether the disparagement can trigger humor. When brands attack or disparage the identifiable target competing brand less severely or more lightly (i.e., low-aggressive humor), the extent to which the normative violation of brand manipulation is perceived by consumers is relatively modest. However, when brands attack or disparage the identifiable target competing brand severely or to extremes (high-aggressive humor), the extent to which the normative violation of brand manipulation is perceived by consumers is relatively excessive. Therefore, compared with excessive norm violations, moderate norm violations are more likely to lead consumers to believe that the brand's disparagement components against the identifiable target competing brand are acceptable and harmless, that is, it improves the benign appraisal of consumers, so it will be more humorous or amusing. Further, there is a link between appreciation of humor and the ability to interact effectively with others (Gervais and David, 2005). Simultaneously perceived humor can lead consumers to like advertisements and brands that initiate teasing in the advertisements (Eisend, 2011). Therefore, compared with high-aggressive humor, low-aggressive humor is more likely to form benign appraisal of consumers, which leads to more appreciation of humor by consumers, which further enhances consumers' interest in the brand and improve consumers' intention to comment and share.

Therefore, we have put forward Hypothesis H2: Benign appraisal plays a mediating effect between low-aggressive humor and consumer engagement behavioral intention; that is, compared with high-aggressive humor, low-aggressive humor generates more benign appraisal, which in turn leads to more consumer engagement behavioral intention.

The moderating effect of brand personality

Benign violation theory also helps explain why similar experiences trigger laughter in some situations but not in others. For example, the same tickling that prompts laughter from a loved one (i.e., considered benign) won't trigger laughter if the tickle is self-inflicted (no violation), nor will it trigger laughter if the tickler is a creepy stranger (i.e., seen as a threat and therefore not playful). In addition, previous research has demonstrated that consumers' evaluation of brand-initiated humorous communication largely depends on the initiator (Romero and Cruthirds, 2006; Warren and McGraw, 2015). Accordingly, we investigate the effects of a boundary condition that seems especially crucial in the current context—that is, the personality of the brand. Brand personality, a set of human characteristics associated with brands, influences how observers perceive brand behavior and its use of humor (Sundar and Noseworthy, 2016). Indeed, it consists of five core dimensions (Aaker, 1997), of which sincerity and excitement are considered the most fundamental (Swaminathan et al., 2009). Sincere brands are viewed as being warm, authentic, consistent, and

family oriented, while exciting brands are associated with youth, uniqueness, fun, and boldness, pushing boundaries (Aaker et al., 2004).

Aaker et al. (2004) pointed out that consumers should react unfavorably toward sincere brands that violate norms excessively. The benign violation theory argued that the playful nature of humorous aggression ads is likely to suggest to consumers that the teaser's behavior is not overly threatening and that the attack is not serious or harmless (McGraw and Warren, 2010; McGraw et al., 2012a). Because sincere brands are associated with goodwill and authenticity, compared with excessive norm-violating high-aggressive humor, when sincere brands use low-aggressive humor, consumers are more likely to interpret the component of brands' attack on the identifiable target competing brand as more appropriate and lightly (Aaker et al., 2004), and therefore is acceptable or harmless, that is, it increases consumers' perception of norm-violating acceptable (i.e., benign appraisal), thus making consumers think that sincere brands are more amusing with low-aggressive humor, ultimately triggering more humor appreciation promotes more consumer engagement behavioral intention. Therefore, when the brand is sincere, low-aggressive humor can enhance consumers' benign appraisal more than high-aggressive humor, and in turn, consumers will be more likely to interact with the brand on social media.

While the use of aggressive humor by exciting brands is somewhat protected from norm violations (Aaker et al., 2004). Therefore, when exciting brands use low-aggressive humor and high-aggressive humor, consumers are more likely to interpret the component of brands' attack on the identifiable target competing brand as not overly threatening, and therefore not serious or harmless (Aaker et al., 2004; McGraw and Warren, 2010). In turn, it is easy to make the consumers think that it is amusing for exciting brands to use low-aggressive humor similar to high-aggressive humor, thus making no significant difference in consumer engagement behavioral intention.

Therefore, we have put forward Hypothesis H3: Brand personality moderates the mediating effect of benign appraisal between low-aggressive humor and consumer engagement behavioral intention; that is, H3a: when the brand is sincere, low-aggressive humor generates more consumers' benign appraisal than high-aggressive humor, which in turn leads to more consumer engagement behavioral intention; H3b: When the brand is exciting, there is no significant difference in consumers' benign appraisal of low-aggressive humor and high-aggressive humor, which in turn makes no significant difference in consumer engagement behavioral intention.

Furthermore, research has shown that when companies use humor on social media to address online public complaints, sincere brands are more likely to use affiliative humor and exciting brands are more likely to use aggressive humor, which in turn is more likely to make online audiences respond more positively (Béal and Grégoire, 2021). Thus, consumers will be more likely to evaluate a brand's dialogue strategy positively when it matches the brand's personality. Specifically, on the one hand, consumers will perceive that the use of less offensive low-aggressive by sincere brands related to intimacy and safety as a better match than the overly offensive high-aggressive humor, so it is easier to accept, which in turn can improve consumers' evaluation of low-aggressive humor, and ultimately enhance consumers' intention to engage. On the other hand, exciting brands are associated with youth and boldness, and audiences would expect exciting brands to be a bit provocative and frivolous (Béal and Grégoire, 2021). Thus, consumers will perceive that the use of low-aggressive humor and high-aggressive humor by exciting brands is matched, so that consumers have no significant difference in evaluation of low-aggressive humor and high-aggressive humor, and ultimately make consumers have no significant difference in consumer engagement behavioral intention.

Therefore, we have put forward Hypothesis H4: Brand personality moderates the effect of low-aggressive humor on consumer engagement behavioral intention; that is, H4a: When the brand is sincere, low-aggressive humor generates more consumer engagement behavioral intention than high-aggressive humor; H4b: When the brand is exciting,

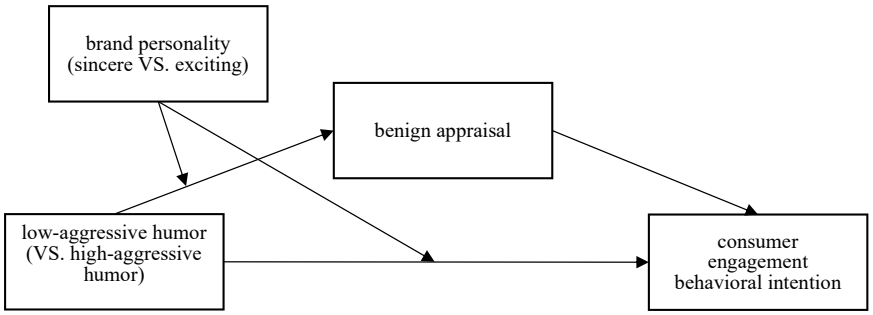


FIGURE 1
Theoretical model.

there is no significant difference in consumer engagement behavioral intention between low-aggressive humor and high-aggressive humor.

Based on the humorous style of brands attacking or disparaging the identifiable target competing brand, this research uses benign violation theory to construct a mediating path of "low-aggressive humor – benign appraisal – consumer engagement behavioral intention"; using brand personality to represent the characteristics of humorous dialogue initiators to explore its moderating effects on the mediating and the dependent variable, thus integrating all research variables to form the theoretical model of this research, as shown in **Figure 1**.

Study 1: Testing the effect of low-aggressive humor (compared with high-aggressive humor) on consumer engagement behavioral intention

This study employs a single-factor (aggressive humor type: low-aggressive humor, low-aggressive humor), between-subjects experiment design to explore that low- aggressive humor (compared with high-aggressive humor) generates more consumer engagement behavioral intention, that is, hypothesis H1.

Study design

Manipulation of aggressive humor type

In order to strengthen the external validity of the research results, this experiment is adapted from real brand attack cases in marketing practice. The experimental stimulus material selected the real mobile phone brand "Samsung." Participants were first informed that Samsung and iPhone are two well-known competing mobile phone brands in China. In order to actively integrate into young consumer groups and better promote new products, Samsung posted on its official Sina Weibo a tweet for dialogue and interaction with its competing brands. Then, by designing different tweets to manipulate aggressive humor type, a total of two experimental materials for experimental conditions were formed. In the low-aggressive humor group, the tweet read: "You're splashing in the bathtub @iphone, I'm taking a dip in the pool, doing what a man should do. Samsung Galaxy S8, IP68 super waterproof system." In the high-aggressive humor group tweeted: "Apple with unreliable waterproof function, no matter how good-looking is, it is also fragile apple @iphone. Samsung specializes in water reversal, and it is more stylish and splash-proof. Samsung Galaxy S8, IP68 super waterproof system."

Variable measurement

The items were all measured on a seven-point Likert scale (1 = strongly disagree at all, 7 = strongly agree), and the detailed measurement items are shown in **Table 1**. Referring to the literature of Pinkleton et al. (2002), three items consisting of ridiculing, mean-spirited, and negative of tweets were used to measure the participants' perceptions of negativity of attack. Referring to the literature on aggressive humor by Speck (1991) and Cline et al. (2003), three items consisting of humorous, teasing, and sarcastic were used to measure participants' perceptions of humor in brand tweets. The measurement of consumer engagement behavioral intention ($\alpha = 0.942$) refers to the mature scale of Pentina et al. (2018). The sentence includes the following six items: I am likely to follow this brand; I am very interested in this brand; I would like this post; I am very likely to comment on this post; I am very likely to repost this post; I am very willing to repost this post. Finally, because the appreciation of humor varies from one individual to another, Warren et al. (2018) emphasized the importance of controlling the sense of humor, which we refer to the scale of Svebak (1996); second, it also controlled for the use of Sina Weibo (yes, no) and brand familiarity and brand favorability.

Participants and study procedure

Based on the purpose of this research is the influence of the brand's aggressive humor style on consumers, the participants in this study are consumers. To conduct study with human participants, this study received ethics approval from the Ethics Review Board of School of Business Administration of Zhongnan University of Economics and Law. All participants signed an informed consent form before they began the experiment.

This research context is social media platform, so the online survey is applicable, and as such, the participants were recruited online. We adopted the non-probabilistic convenience sampling method to collect data through professional online survey sites[1]. Respondents were provided CNY0.50 for their participation at the end of the survey.

Participants were randomly assigned to one of two experimental conditions and then read a tweet from a brand to promote a new product, and then asked participants to rate the brand's engagement behavioral intention in the read tweets, and to provide their perceptions of negativity and perceptions of humor in brand dialogue. It also includes filling in brand familiarity, brand favorability, sense of humor, Sina Weibo usage, and relevant demographic information. In order to test whether the participants were attentive in the process of filling out the questionnaire, this experiment designed an attention test question. The item is "What brand did you evaluate in

1 wjx.com

TABLE 1 Questionnaire items.

Construct	Item code	Item
Perceptions of negativity of attack	PNA1	I think this tweet is ridiculing
	PNA2	I think this tweet is mean-spirited
	PNA3	I think this tweet is negative
Perceptions of humor	PH1	I think this tweet is humorous
	PH2	I think this tweet is teasing
	PH3	I think this tweet is sarcastic
Perceptions of sincere	PS1	The brand could be described as sincere
	PS2	The brand could be described as warm
	PS3	The brand could be described as wholesome
	PS4	The brand could be described as family oriented
Perception of exciting	PE1	The brand could be described as exciting
	PE2	The brand could be described as unique
	PE3	The brand could be described as young
	PE4	The brand could be described as trendy
Benign appraisal	BA1	I think this tweet is not threatening
	BA2	I think this tweet is generally not a big problem
	BA3	I think this tweet is acceptable
	BA4	I think this tweet is reasonable
Perceptions of usefulness	PU1	This tweet helped me understand the product
	PU2	This tweet helped me with my purchase decision
	PU3	The tweet contained important product information
	PU4	This tweet provided me with useful product information
Perceptions of coolness	PC1	I think this tweet is cool
	PC2	I think this tweet is hip
	PC3	I think this tweet is trendy
	PC4	I think this tweet is stylish
Sense of humor	SH1	I easily recognize a hint, such as a wink or a slight change in emphasis, as a mark of humorous intent
	SH2	It is easy for me to find something comical, witty, or humorous in most situations
	SH3	I have much cause for amusement during an ordinary day
Use of Sina Weibo	USW1	Do you have a Sina Weibo account
Brand familiarity	BF1	How familiar are you with Samsung's mobile phone brand
Brand favorability	BV1	How do you like Samsung's mobile phone brand
Consumer engagement behavioral intention	CE1	I am likely to follow this brand
	CE2	I am very interested in this brand
	CE3	I would like this post
	CE4	I am very likely to comment on this post
	CE5	I am very likely to repost this post
	CE6	I am very willing to repost this post

this survey?," and the options are A. Lenovo, B. Coca-Cola, C. Samsung, D. BMW. If the participant's choice is not option C, it is reasonable to think that the participant did not fill in the questionnaire carefully according to the experimental requirements, and it will be excluded as an invalid questionnaire. Finally, this experiment also designed a hypothesis guessing question. The item is "What is the hypothesis of this study?" In addition, we controlled the IP address to ensure that each volunteer participant only answered the questionnaire once. After data collection, we conducted strict screening and deleted the questionnaires with obvious regularity and too short response time. Among them, according to the average fastest reading of 500 words per minute by most people, the cases whose response time is less than 31 s were excluded.

Study results

Manipulation testing

Excluding 32 people who failed the attention check, 46 people who guessed the hypothesis, and 112 people who answered the questions with obvious regularity and too short response time, the final sample was 620, of which 66.6% were male and 33.4% were female. The demographic breakdown of the sample is shown in Table 2 and the descriptive statistics for key constructs are shown in Table 3. All experimental measurement items loadings, AVE, and CR are included in Table 4. All experimental measurement model indices are included in Table 5. The results in Table 4 show that the measurement item loadings of all constructs exceed the reference value of 0.5, which supports convergent validity; The average variance extracted (AVE) values of all constructs exceeded the cutoff value of 0.50 and the AVE square root values are greater than inter-construct correlations (as shown in Table 6), which supports discriminant validity (Fornell and Larcker, 1981).

The aggressive humor type manipulation test was successful. (a) Taking two experimental conditions (i.e., aggressive humor type) as the independent variable, and perceptions of negativity as the dependent variable, conduct paired-samples t-tests. The results showed that compared with the high-aggressive humor group (M = 5.353, SD = 0.928), the low-aggressive humor group (M = 3.082, SD = 1.533) perceived less negativity of attack [$t(309) = -23.152$, $P < 0.01$]; (b) The results of one-tailed t-tests respectively showed that the humor level of the low-aggressive humor group [M = 4.424, SD = 0.945, $t(309) = 7.891$, $P < 0.01$] and the high-aggressive humor group [M = 4.586, SD = 0.839, $t(309) = 12.294$, $P < 0.01$] was significantly greater than the mid-point (4) of the seven-point scale measuring humor.

Hypothesis testing

Taking aggressive humor type as the independent variable and the potential covariates as the dependent variable, a one-way MANOVA analysis was conducted. The results showed that there was a significant difference in brand familiarity [$F(1,618) = 5.666$, $P = 0.018$], brand favorability [$F(1,618) = 16.191$, $P < 0.01$], Sina Weibo usage

TABLE 2 Demographic breakdown.

Variable		Study 1		Study 2		Study 3	
		N	(%)	*N*	(%)	*N*	(%)
Gender	Male	413	66.6	317	51.1	178	55.6
	Female	207	33.4	303	48.9	142	44.4
	Total	620	100	620	100	320	100
Age	19 years old or below	18	2.9	33	5.3	13	4.1
	20–29 years old	359	57.9	356	57.4	194	60.6
	30–39 years old	209	33.7	211	34.0	97	30.3
	40–49 years old	31	5.0	17	2.7	12	3.8
	50 years old or above	3	0.5	3	0.5	4	1.3
	Total	620	100	620	100	320	100
Education background	high school or below	70	11.3	55	8.9	30	9.4
	College or bachelor	460	74.2	478	77.1	252	78.8
	Master or above	90	14.5	87	14.0	38	11.9
	Total	620	100	620	100	320	100
Use of Sina Weibo	Yes	590	95.2	596	96.1	310	96.9
	No	30	4.8	24	3.9	10	3.1
	Total	620	100	620	100	320	100

[$F(1,618) = 5.069$, $P = 0.025$] and gender [$F(1,618) = 7.974$, $P < 0.01$] in the two aggressive humor types, and there was no significant difference in sense of humor, age and education background in the two aggressive humor types (all Ps > 0.10). Then, taking aggressive humor type as the independent variable, consumer engagement behavioral intention as the dependent variable, and sense of humor, Weibo usage, brand familiarity, brand favorability, and demographic variables as the control variables, an ANCOVA analysis was conducted. The results showed that there was a significant difference in consumer engagement behavioral intention in the two aggressive humor types [$F(1,611) = 17.426$, $P < 0.01$]. Low-aggressive humor (M = 5.441, SD = 1.187) obtained more consumer engagement behavioral intention than high-aggressive humor (M = 4.843, SD = 1.416). Therefore, hypothesis H1 is verified.

Study 1 preliminarily verified that the low-aggressive humor group had a stronger engagement behavioral intention in social media than the high-aggressive humor control group. On this basis, the next study 2 will use brands of different product types to conduct experiments to further consolidate the results of study 1 and further explore the mediating mechanism of low-aggressive humor affecting consumer engagement behavioral intention.

Study 2: Testing the mediating role of benign appraisal between low-aggressive humor and consumer engagement behavioral intention

On the basis of study 1, this study adopts a between-subjects experiment (low-aggressive humor vs. high-aggressive humor), the main purpose is to examine the mediating effect of benign appraisal between low-aggressive humor and consumer engagement behavioral intention, that is H2, thereby explaining why low-aggressive humor is more effective than high-aggressive humor. This study also re-examines the relative advantage of low-aggressive humor, by comparing the effect of low- aggressive humor and high-aggressive humor on consumer engagement behavioral intention, that is H1.

Study design

The fictitious cake brand "TastyBakes" was chosen for the experimental stimulus material to avoid the interference of

TABLE 3 Descriptive statistics.

Construct	Study 1		Study 2		Study 3	
	Low-aggressive humor	High-aggressive humor	Low-aggressive humor	High-aggressive humor	Low-aggressive humor	High-aggressive humor
Benign appraisal	–	–	M = 5.102	M = 4.522	M = 5.163	M = 4.470
	–	–	SD = 1.178	SD = 1.269	SD = 0.944	SD = 1.359
Consumer engagement behavioral intention	M = 5.441	M = 4.843	M = 5.054	M = 4.472	M = 5.174	M = 4.495
	SD = 1.187	SD = 1.416	SD = 1.454	SD = 1.222	SD = 1.091	SD = 1.364
Sense of humor	M = 5.556	M = 5.447	M = 5.298	M = 5.075	M = 5.404	M = 5.175
	SD = 0.978	SD = 0.968	SD = 1.088	SD = 1.046	SD = 0.915	SD = 1.074
Brand familiarity	M = 5.620	M = 5.400	–	–	–	–
	SD = 1.108	SD = 1.185	–	–	–	–
Brand favorability	M = 5.450	M = 5.030	–	–	–	–
	SD = 1.213	SD = 1.377	–	–	–	–

10.3389/fpsyg.2022.966254

TABLE 4 Measurement model.

Construct	Item code	Item loadings	AVE	CR
Perceptions of negativity of attack				
	PNA1	0.803[a]/0.844[b]/0.874[c]	0.781[a]/0.712[b]/0.768[c]	0.914[a]/0.881[b]/0.908[c]
	PNA2	0.928[a]/0.797[b]/0.928[c]		
	PNA3	0.914[a]/0.888[b]/0.823[c]		
Perceptions of sincere				
	PS1	0.802[c]	0.504[c]	0.800[c]
	PS2	0.765[c]		
	PS3	0.640[c]		
	PS4	0.614[c]		
Perception of exciting				
	PE1	0.642[c]	0.564[c]	0.837[c]
	PE2	0.734[c]		
	PE3	0.794[c]		
	PE4	0.821[c]		
Benign appraisal				
	BA1	0.782[b]/0.721[c]	0.658[b]/0.623[c]	0.884[b]/0.867[c]
	BA2	0.746[b]/0.662[c]		
	BA3	0.861[b]/0.911[c]		
	BA4	0.849[b]/0.838[c]		
Sense of humor				
	SH1	0.782[a]/0.849[b]/0.760[c]	0.579[a]/0.614[b]/0.555[c]	0.805[a]/0.826[b]/0.789[c]
	SH2	0.710[a]/0.750[b]/0.737[c]		
	SH3	0.789[a]/0.747[b]/0.737[c]		
Consumer engagement behavioral intention				
	CE1	0.807[a]/0.806[b]/0.728[c]	0.728[a]/0.701[b]/0.661[c]	0.941[a]/0.933[b]/0.921[c]
	CE2	0.811[a]/0.780[b]/0.751[c]		
	CE3	0.855[a]/0.889[b]/0.844[c]		
	CE4	0.838[a]/0.794[b]/0.811[c]		
	CE5	0.905[a]/0.866[b]/0.877[c]		
	CE6	0.899[a]/0.881[b]/0.855[c]		

[a]Study 1 results.
[b]Study 2 results.
[c]Study 3 results.

existing brands in the market and their brand image on the experimental results.

Manipulation of aggressive humor type

Referring to the experiment of Thomas and Fowler (2021), the experimental material is realized through the design of brand tweets. Participants were asked to imagine seeing a tweet from a consumer on Sina Weibo that read: "Just arrived abroad and need to buy a birthday cake for the party. How does @TopFrost compare to @TastyBakes?" TastyBakes and TopFrost are two competing cake brands in foreign countries. In order to better dialogue with consumers, promote consumers' understanding of the brand and its products. TastyBakes responded to the consumer on its official Weibo account. We manipulated the aggressive humor type by designing different tweets, aggregated to form experimental material for two experimental conditions. In the low-aggressive humor group, the tweet read: "If you want frosting (in your tweets or in your cake), go to @TopFrost; if you want tasty food, go to @TastyBakes." In the high aggressive humor group, the tweet read: "@TastyBakes food, always tasty, never frosty, please avoid @TopFrost."

Variable measurement

The items were all measured on a seven-point Likert scale (1 = strongly disagree at all, 7 = strongly agree), and the detailed measurement items are shown in **Table 1**. The measures of perceptions of negativity of attack, perceptions of humor, and consumer engagement behavioral intention (α = 0.933) are consistent with study 1. The measurement of benign appraisal (α = 0.884) is adapted from previous studies by scholars (McGraw and Warren, 2010; McGraw et al., 2015; Warren and McGraw, 2015), and the measurement includes the following four items: I think this tweet is not threatening; I think this tweet is generally not a big problem; I think this tweet is acceptable;

I think this tweet is reasonable. In addition, since previous research found that both information search (Java et al., 2007) and coolness attention (Toubia and Stephen, 2013) can motivate consumer engagement on Twitter, the two items of perceptions of usefulness and perceptions of coolness were added to exclude alternative explanations of consumer engagement behavioral intention. The perceptions of usefulness ($\alpha = 0.911$) refers to the scale of Park and Lee (2009), which consists of four items. The perceptions of coolness ($\alpha = 0.929$) refers to the scale of Rahman (2013), which consists of four items. Finally, sense of humor and Sina Weibo usage were also controlled.

Participants and study procedure

As the participants in this study are the same consumers as those in study 1, the ethics review was also conducted in this study and the informed consent was provided for the participants. As this study is an online survey as in study 1, this study also recruited participants through online survey sites (i.e., see text footnote 1). Respondents were provided CNY1.00 for their participation.

Manipulating the aggressive humor type, a questionnaire was pre-tested on 38 participants. After the pre-test was passed and the formal experiment was conducted, participants were randomly assigned to one of two experimental conditions and read a tweet from the brand in the experimental condition. Participants were then asked to rate the brand's engagement behavioral intention in the read tweets, and to provide perceptions of negativity of attack and perceptions of humor in brand dialogue, and then to provide benign appraisal, perceptions of usefulness, perceptions of coolness. Also included are sense of humor, Sina Weibo usage, and relevant demographic information. Finally, the participants filled out the attention test question and the hypothesis guessing question. The attention test question is "What type of product does the brand in the reading material belong to?," and the options are A. Car, B. Mobile phone, C. Food, D. Clothing. The hypothesis guessing question is the same as in study 1. In addition, this study controlled the IP address as in study 1. After data collection, consistent with study 1, this study deleted the questionnaires with obvious regularity and too short response time.

Study results

Manipulation testing

Excluding 38 people who failed the attention check, 30 people who guessed the hypothesis, and 152 people who answered the questions with obvious regularity and too short response time, the final sample was 620, of which 51.1% were male and 48.9% were female. The demographic breakdown of the sample is shown in Table 2 and the descriptive statistics for key constructs are shown in Table 3. All experimental measurement items loadings, AVE, and CR are included in Table 4. The results in Tables 4–6 support the convergent and discriminant validity of all constructs.

The aggressive humor type manipulation test was successful. (a) The results of paired-samples t-tests showed that compared with the high-aggressive humor group ($M = 4.846$, $SD = 1.022$), the low-aggressive humor group ($M = 3.475$, $SD = 1.622$) perceived less negativity of attack [$t(309) = -12.079$, $P < 0.01$]; (b) The results of one-tailed t-tests respectively showed that the humor level of the low-aggressive humor group [$M = 4.370$, $SD = 1.311$, $t(309) = 4.968$, $P < 0.01$] and the high-aggressive humor group [$M = 4.559$, $SD = 1.057$, $t(309) = 9.310$, $P < 0.01$] was significantly greater than the mid-point (4) of the seven-point scale measuring humor.

Hypothesis testing

Taking aggressive humor type as the independent variable and the potential covariates as the dependent variable, a one-way MANOVA analysis was conducted. The results showed that there was a significant difference in sense of humor [$F(1,618) = 6.741$, $P = 0.010$] and gender [$F(1,618) = 6.246$, $P = 0.013$] and education background [$F(1,618) = 5.619$, $P = 0.018$] in the two aggressive humor types, and there was no significant difference in Sina Weibo usage and age in the two aggressive humor types (all Ps > 0.10). Then, ① with aggressive humor type as the independent variable, consumer engagement behavioral intention as the dependent variable, and sense of humor, Weibo usage, and demographic variables as the control variables, an ANCOVA analysis was conducted. The results showed that there was a significant difference in consumer engagement behavioral intention in the two aggressive humor types [$F(1,613) = 21.508$, $P < 0.01$], and the low-aggressive humor ($M = 5.054$, $SD = 1.454$) obtained more consumer engagement behavioral intention than high-aggressive humor ($M = 4.472$, $SD = 1.222$). Therefore, H1 is verified again. ② With aggressive humor type as the independent variable, benign appraisal as the dependent variable, sense of humor, Weibo usage, and demographic variables as the control variables, an ANCOVA analysis was conducted. The results showed that there was a significant difference in benign appraisal in the two aggressive humor types [$F(1,613) = 26.746$, $P < 0.01$], and the low-aggressive humor ($M = 5.102$, $SD = 1.178$) obtained more benign appraisal than high-aggressive humor ($M = 4.522$, $SD = 1.269$). ③ With aggressive humor type as the independent variable, perceptions of usefulness and perceptions of coolness as the dependent variable, sense of humor, Weibo usage, and demographic variables as the control variables, an ANCOVA analysis was conducted. The results showed that there was no significant difference in perceptions of usefulness in the two aggressive humor types [$F(1,613) = 2.667$, $P = 0.103$]. There was no significant difference in perceptions of coolness in the two aggressive humor types

TABLE 5 Measurement model indices.

Studies	Factors	Items	CFI	TLI	RASEA	SRMR	χ^2	d.f.	*p*
2	4	16	0.947	0.934	0.079	0.050	468.89	96	<0.001
3	6	24	0.916	0.900	0.073	0.055	632.40	233	<0.001

[$F(1,613) = 2.414$, $P = 0.121$]. Thus, alternative explanations for perceptions of usefulness and perceptions of coolness are excluded.

Mediation effect testing

To test the mediating effect of benign appraisal between low-aggressive humor and consumer engagement behavioral intention, the aggressive humor type was set as a dummy variable (high-aggressive humor was coded as 0, low-aggressive humor was coded as 1), and sense of humor, Weibo usage, and demographic variables were used as the control variables. The maximum variance inflation factor (VIF) value in the models were 1.665, which was less than 3.3, indicating that variables did not contain problematic collinearity (Petter et al., 2007). As shown in the results of process model 4 (**Figure 2**), the mediating effect of benign appraisal was significant (indirect effect was β = 0.190, SE = 0.043; 95% CI = [0.091, 0.312]), that is, compared with high-aggressive humor, low-aggressive humor promoted consumer engagement behavioral intention by generating higher benign appraisal, which validates H2.

Study 3: Testing the moderating effect of brand personality

This study is a 2 (aggressive humor type: low-aggressive vs. high-aggressive humor) × 2 (brand personality type: sincere vs. exciting) between-subjects experiment, the main purpose is to examine that brand personality moderates low-aggressive humor on consumer engagement behavioral intention, that is H3, and to examine brand personality moderates the mediating effect of benign appraisal between low-aggressive humor and consumer engagement behavioral intention, that is H4.

TABLE 6A Discriminant validity test of variables in study 1.

	1	2	3
1. Perceptions of negativity of attack	**0.884**		
2. Sense of humor	0.020	**0.761**	
3. Consumer engagement behavioral intention	−0.193**	0.468**	**0.853**

$n = 620$.
**$p < 0.01$.
The data marked in bold in the diagonal line of the matrix is the AVE square root, whereas the rest of the data is the corresponding correlation.

Study design

Manipulation of aggressive humor type and brand personality type

Manipulating aggressive humor type through different tweets, manipulating brand personality types through different brands' official Weibo banner images, company profiles, and brand logos and tweets, forming experimental materials for four experimental conditions in total. Following the procedure recommended by Aaker et al. (2004), first, brand personality was manipulated through vocabulary choice and phrasing. Use "hello" and "opponents" for sincere brands, and "hey" and "friends" for exciting brands. Use the hashtags "#family," "#warm" and "#wholesome" for sincere brands and "#unique," "#young" and "#trendy" for exciting brands. Second, the brand personality is manipulated by manipulating the banner image, company profile, and brand logo of the brand's official Weibo page. Use an image of families for sincere brands, and an image of jumping young people for exciting brands. The company profile manipulation for sincere brand is: "Dear customer, we sincerely welcome you to Bigmeat. We are a #family#warm#wholesome fast food brand." The company profile manipulation for exciting brands is: "Hey guy, we enthusiastically welcome you to Topburger. We are

TABLE 6B Discriminant validity test of variables in study 2.

	1	2	3	4
1. Perceptions of negativity of attack	**0.844**			
2. Benign appraisal	−0.115**	**0.811**		
3. Sense of humor	0.034	0.566**	**0.784**	
4. Consumer engagement behavioral intention	−0.154**	0.633**	0.533**	**0.837**

$n = 620$.
**$p < 0.01$.
The data marked in bold in the diagonal line of the matrix is the AVE square root, whereas the rest of the data is the corresponding correlation.

TABLE 6C Discriminant validity test of variables in study 3.

	1	2	3	4	5	6
1. Perceptions of negativity of attack	**0.876**					
2. Perceptions of sincere	0.083	**0.710**				
3. Perception of exciting	−0.099	0.244**	**0.751**			
4. Benign appraisal	−0.331**	0.045	0.338**	**0.789**		
5. Sense of humor	0.040	0.344**	0.486**	0.327**	**0.745**	
6. Consumer engagement behavioral intention	−0.203**	0.191**	0.435**	0.667**	0.308**	**0.813**

$n = 320$.
**$p < 0.01$.
The data marked in bold in the diagonal line of the matrix is the AVE square root, whereas the rest of the data is the corresponding correlation.

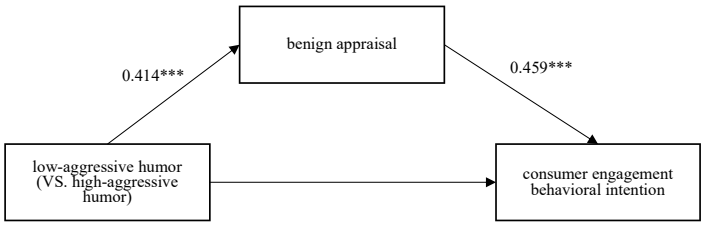


FIGURE 2
Path coefficient of low-aggressive humor on consumer engagement behavioral intention. Significance levels: $P < 0.001$ (***).

a #unique#young#trendy fast food brand." Using blue font, the brand logo of Comic Sans font represents sincere brands; using orange font, the brand logo of **phosphate** font represents exciting brands. Therefore ① low-aggressive × exciting group: participants were asked to read a tweet of Topburger' s aggressive humor to Bigmeat on Weibo that read: "Hey! @Bigmeat your cheeseburgers' beef is freezing; our burgers' beef is so fresh and too cool to ever be frozen. –From your friend @Topburger" ② low-aggressive × sincere group: participants were asked to read a tweet of Bigmeat' s aggressive humor to Topburger on Weibo that read: "Hello! @Topburger your cheeseburgers' beef is freezing; our burgers' beef is so fresh and too cool to ever be frozen. – From your opponent @Bigmeat" ③ high-aggressive × exciting group: participants were asked to read a tweet of Topburger' s aggressive humor to Bigmeat on Weibo that read: "Hey! Never thought your cheeseburgers are full of frozen beef @Bigmeat; our burgers' beef is so fresh and so cool that it never be frozen. – From your friend @Topburger" ④ high-aggressive × sincere group: participants were asked to read a tweet of Bigmeat' s aggressive humor to Topburger on Weibo that read: "Hello! Never thought your cheeseburgers are full of frozen beef @Topburger; our burgers' beef is so fresh and so cool that it never frozen. – From your opponent @Bigmeat"

Variable measurement

The items were all measured on a seven-point Likert scale (1 = strongly disagree at all, 7 = strongly agree), and the detailed measurement items are shown in Table 1. The measures of perceptions of negativity of attack, perceptions of humor, benign appraisal (α = 0.865), and consumer engagement behavioral intention (α = 0.920) are consistent with study 2. Referring to the literature of Aaker et al. (2004) on perceptions of sincere and perception of exciting, four items consisting of sincere, warm, wholesome, and family oriented were used to measure participants' perceptions of sincere; four items consisting of exciting, unique, young, and trendy were used to measure the participants' perceptions of exciting.

Participants and study procedure

As the participants in this study are the same consumers as those in study 1, the ethics review was also conducted in this study and the informed consent was provided for the participants. As this study is an online survey as in study 1, this study also recruited participants through online survey sites (i.e., see text footnote 1). Respondents were provided CNY1.50 for their participation.

Participants were randomly assigned to one of the four experimental conditions, then read the company profile of the brand's official Weibo page in the experimental condition, and then asked participants to provide perceptions of sincere and perceptions of exciting. Then read a tweet from the brand that had a dialogue and interaction with its competing brands for promotion, and then asked the participants to rate the brand's engagement behavioral intention in the read tweets, and to provide perceptions of negativity of attack and perceptions of humor in brand dialogue, and fill in appraisal. Also included are sense of humor, Sina Weibo usage, and relevant demographic information. Finally, the participants filled in the attention test question and the hypothesis guessing question, which are the same as in study 2. In addition, this study controlled the IP address as in study 1. After data collection, consistent with study 1, this study deleted the questionnaires with obvious regularity and too short response time.

Study results

Manipulation testing

Excluding 26 people who failed the attention check, 16 people who guessed the hypothesis, and 88 people who answered the questions with obvious regularity and too short response time, the final sample was 320, of which 55.6% were male and 44.4% were female. The demographic breakdown of the sample is shown in Table 2 and the descriptive statistics for key constructs are shown in Table 3. All experimental measurement items loadings, AVE, and CR are included in Table 4. The results in Tables 4–6 support the convergent and discriminant validity of all constructs.

The manipulation test was successful. ① (a) The results of between-subjects ANOVA showed that compared with the high-aggressive humor group (M = 4.235, SD = 1.535), the low-aggressive humor group (M = 3.590, SD = 1.674) perceived less negativity of attack [$F(1,316) = 13.156$, $P < 0.01$]; (b) The brand personality or the interaction between aggressive

humor and brand personality did not have a significant effect on perceptions of negativity of attack (all *P*s > 0.10). ② The results of one-tailed *t*-tests respectively showed that the humor level of the low-aggressive humor group [M = 4.279, SD = 1.200, $t(159) = 2.943$, $P < 0.01$] and the high-aggressive humor group [M = 4.494, SD = 1.118, $t(159) = 5.587$, $P < 0.01$] was significantly greater than the mid-point (4) of the seven-point scale measuring humor. ③ The perceptions of sincere in the sincere brand group (M = 5.381, SD = 0.964) was significantly higher [$F(1,316) = 43.908$, $P < 0.01$] than that of the exciting brand group (M = 4.700, SD = 0.888). ④ The perceptions of exciting in the exciting brand group (M = 5.744, SD = 0.751) was significantly higher [$F(1,316) = 124.913$, $P < 0.01$] than that of the sincere brand group (M = 4.678, SD = 0.943). ⑤ The aggressive humor or the interaction between aggressive humor and brand personality did not have a significant effect on perceptions of brand personality (all *P*s > 0.10). ⑥ There was no correlation between the perceptions of sincere scale and the perceptions of exciting scale ($r = -0.02$, $P > 0.10$), indicating that these two brand personalities are exclusive, distinct, and independent (Aaker, 1997).

Hypothesis testing

Taking aggressive humor type and brand personality as the independent variable and the potential covariates as the dependent variable, a two-way MANOVA analysis was conducted. The results showed that there was no significant difference in sense of humor, Sina Weibo, and all demographic variables in four experimental conditions (all *P*s > 0.10). Then, ① with aggressive humor type and brand personality as the independent variables, and consumer engagement behavioral intention as the dependent variable, an ANOVA analysis was conducted. The results showed that (a) aggressive humor type had a significant effect on consumer engagement behavioral intention [$F(1,316) = 26.823$, $P < 0.01$]; (b) brand personality had a significant effect on consumer engagement behavioral intention [$F(1,316) = 27.988$, $P < 0.01$]; (c) Importantly, the interaction between aggressive humor and brand personality had a significant effect on consumer engagement behavioral intention [$F(1,316) = 8.545$, $P < 0.01$]. The simple effect analysis further showed (**Figure 3**) that in the sincere brand condition, low-aggressive humor (M = 5.019, SD = 1.134) generated more consumer engagement behavioral intention [$F(1,316) = 32.824$, $P < 0.01$] than high-aggressive humor (M = 3.956, SD = 1.311), which verified H4a; in the exciting brand condition, low-aggressive humor (M = 5.329, SD = 1.030) and high-aggressive humor (M = 5.033, SD = 1.200) was no significant difference in consumer engagement behavioral intention [$F(1,316) = 2.545$, $P = 0.112$], which verified H4b. ② Similarly, with aggressive humor type and brand personality as the independent variables, and benign appraisal as the dependent variable, an ANOVA analysis was conducted. The results showed that (a) aggressive humor type had a significant effect on benign appraisal [$F(1,316) = 31.693$, $P < 0.01$]; (b) brand personality had a significant effect on benign appraisal [$F(1,316) = 35.220$, $P < 0.01$]; (c) the interaction between aggressive humor type and brand personality had a significant effect on benign appraisal [$F(1,316) = 8.617$, $P < 0.01$]. The simple effect analysis further showed that in the sincere brand condition, low-aggressive humor (M = 4.978, SD = 1.005) generated more benign appraisal [$F(1,316) = 36.681$, $P < 0.01$] than high-aggressive humor (M = 3.925, SD = 1.312); in the exciting brand condition, low-aggressive humor (M = 5.347, SD = 0.845) and high-aggressive humor (M = 5.016, SD = 1.180) was no significant difference in benign appraisal [$F(1,316) = 2.629$, $P > 0.01$].

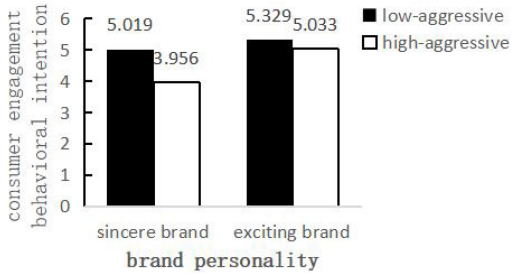


FIGURE 3
The influence of aggressive humor type and brand personality on consumer engagement behavioral intention.

Mediation effect testing

Test the significance of the moderated-mediation model involving aggressive humor type as the independent variable, benign appraisal as the mediator variable, brand personality as the moderator variable, consumer engagement behavioral intention as the dependent variable, demographic variables, sense of humor, and Sina Weibo usage as control variables. The maximum variance inflation factor (VIF) value in the models were 3.123, which was less than 3.3, indicating that variables did not contain problematic collinearity (Petter et al., 2007). The process model 8 (moderated-mediation model) results showed that the mediating effect of benign appraisal was significant (β = 0.462, SE = 0.147; 95% CI [0.179, 0.756]). This indicates that the indirect effect of "aggressive humor type-benign appraisal-consumer engagement behavioral intention" is different depending on the brand personality conditions; specifically, this given sequence is significant for sincere brands (β = 0.582, SE = 0.130; 95%CI [0.336, 0.849]), which verified H3a; but not significant for exciting brands (β = 0.120, SE = 0.090; 95% CI [−0.053, 0.301]), which verified H3b.

Discussion of the results

Through study 1 with mobile phone brands, we found that low-aggressive humor promotes consumer engagement

behavioral intention more than high-aggressive humor. This conclusion is consistent with the research conclusions of scholars such as Thomas and Fowler (2021). On this basis, study 2 using cake brands to examine that low-aggressive humor can promote consumer engagement behavioral intention; benign appraisal play a mediating role between low-aggressive humor and consumer engagement behavioral intention. Distinguished from the study of Thomas and Fowler (2021), we explore the impact of low-aggressive humor from a positive cognitive perspective (benign appraisal) as opposed to negative cognitive factors (perceptions of manipulative intent). This result demonstrates the relationship between aggressive humor and consumer engagement behavioral intention from a new perspective. In addition, study 3 with the hamburger brand further compared the effect of two different aggressive humor on consumer engagement behavioral intention in two different brand personalities. It was found that sincere brands used low-aggressive humor to promote consumer engagement behavioral intention more than high-aggressive humor, consistent with the conclusion of previous study. Béal and Grégoire (2021) have proved that brand personality moderated the effect of humor type and likes/retweets/purchase intentions. Further, brand personality moderates the mediating effect of benign appraisal between low-aggressive humor and consumer engagement behavioral intention. Specifically, sincere brands used low-aggressive humor to increase consumers' benign appraisal more than high-aggressive humor, and then it is easier to enhance the consumer engagement behavioral intention. This result is similar to that of the study of Béal and Grégoire (2021), that is, brand personality moderates the mediating effect of humor appreciation between humor type and likes/retweets/purchase intentions. These two results enrich the boundary condition of the relationship between brand-to-brand teasing and consumer engagement behavioral intention.

Theoretical contributions and future research

Theoretical contributions

(1) This research has enriched the connotation of humorous advertisements and the research on the influence effect of the interaction between brands. Although the impact of traditional non-brand anthropomorphic humor advertisements on consumers has been extensively studied, little research has examined the influence of brand anthropomorphic humorous advertisements on consumers (Eisend, 2009; Nair, 2016; Warren et al., 2019). This research enriches the connotation of humorous advertisements by exploring the influence of brand-to-brand teasing on consumer engagement behavioral intention. Moreover, existing research has mainly focused on how brand-consumer interaction affects consumers' cognition (MacInnis and Folkes, 2017), but has not paid particular attention to how brand-brand interaction affects consumers' cognition. This research extends the theory of interpersonal humor to the field of brand dialogue strategies, thereby dividing the aggressive dialogue styles that occur between brands on social media into two forms of humor, exploring the styles of brands' aggressive dialogue in social media, and its influence on consumers' cognition and behavioral intentions, which enriches the connotation of humorous advertisements and the research on the influence effect of the interaction between brands.

(2) This research has expanded the mechanism of the positive effect of low-aggressive humor between brands. A small number of existing studies on low-aggressive humor between brands have explored the effect of low-aggressive humor on brand interest through the role of negative cognitive factors such as perceptions of manipulative intent (Thomas and Fowler, 2021), while ignoring the role of positive evaluation factors. Based on benign violation theory, this research uses benign appraisal as a mediator to explore the mechanism of low-aggressive humor in promoting consumer engagement behavior from the perspective of positive evaluation, and enriches the application of benign violation theory in the field of inter-brand interaction research.

(3) This research has enriched the theory of brand personality. Although previous research on low-aggressive humor between brands focus on the moderating effect of the object brand factor of dialogue interaction (Thomas and Fowler, 2021), there has been very little examination of the moderating effect of humor initiator factors. Based on benign violation theory, consumers' evaluation of brand-initiated humorous communication largely depends on the initiator (Romero and Cruthirds, 2006; Warren and McGraw, 2015; Béal and Grégoire, 2021). Therefore, from the perspective of the characteristics of the subject brand itself in dialogue interaction, this research explores the moderating effect of brand personality on the relationship between the brand's low-aggressive humor and consumers' engagement behavioral intention, thereby expanding the boundary condition that low-aggressive humor between brands brings positive effects also enriches the theory of brand personality.

Managerial implications

Based on the social media context, this research studies the influence mechanism and boundary condition of brand-to-brand aggressive humor style on consumers' engagement behavioral intention. The results of the study will bring inspiration to brands in terms of brand impression management, brand communication, and consumer behavior in the effective use of aggressive humor style on social media.

(1) On social media, brands can expand from focusing only on the brand-to-consumer relationship to also focusing on the brand-to-brand relationship. Even brands shouldn't be afraid to engage in aggressive dialogue interaction with other brands, as brands creating aggressive humor dialogue with another brand can promote consumer engagement behavior. In particular, when brands use low-aggressive humor, it is more popular than high-aggressive humor. Therefore, marketers can actively use brand-teasing content strategies on social media brand pages to attract users on social media, thereby evoking user interaction with the brand.

(2) The dialogue between brands shapes consumers' evaluation of brands and consumers' engagement behavioral intention in social media. Brands should create content that obtains consumers' benign appraisal. In particular, when a brand initiates an aggressive humor dialogue with another brand and uses low-aggressive humor, it can generate higher benign appraisal than high-aggressive humor. Therefore, brands should actively create low-aggressive humorous content that generates more benign appraisal.

(3) In social media context, brand teasing must also be in line with the brand's personality in order to obtain consumers' benign appraisal, and the brand should adjust its humor style according to the brand's personality type. Exciting brands are more likely to use aggressive humor than sincere brands. Especially, compared with low-aggressive humor, exciting brands can also achieve good brand communication effects by using high-aggressive humor.

Research limitations and future research

First, this research only collects humorous dialogues between mobile phone brands and retail food brands on social media, and future research needs to be extended to brands of other product types to improve the generalizability of the results. Second, this research only studies the humorous attacks that brands selectively target specific target brands, and frequently attacking or attacking too many target brands may damage the evaluation of the attacking brand (Kowalski, 2004), and moderators such as the frequency of brand attacks and the number of target brands of brand attacks need to be included for further discussion. Third, this research only selects brand dialogues on Sina Weibo social media, but the social norms of different social media are different. In the future, we can study whether the marketing effect brought by the aggressive humor style of other social media is consistent with the conclusions of this research. Fourth, due to the different understandings of humor in different cultures, humor tolerance, and normative beliefs, this research only considers the marketing effect brought by the brand aggressive humor style based on the Chinese context. In the future, we can explore the marketing effects of brand aggressive humor style on Western cultural backgrounds, individuals with different humor tolerance and individuals with different normative beliefs.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving human participants were reviewed and approved by the Ethics Review Board of School of Business Administration of Zhongnan University of Economics and Law. All participants signed an informed consent form. The patients/participants provided their written informed consent to participate in this study.

Author contributions

Y-MN wrote and revised the manuscript. CH and T-TT collected and discussed the results. DL analyzed the data and revised the manuscript. All authors contributed to the article and approved the submitted version.

Funding

The work was financially supported by Higher Education Undergraduate Teaching Reform Project of Yulin Normal University (Project No. 2021XJJGYB17).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Aaker, J. L. (1997). Dimensions of brand personality. *J. Mark. Res.* 34, 347–356. doi: 10.1177/002224379703400304

Aaker, J., Fournier, S., and Brasel, S. A. (2004). When good brands do bad. *J. Consum. Res.* 31, 1–16. doi: 10.1086/383419

Apter, M.J. (1982). The experience of motivation: The theory of psychological reversals. London: Academic Press.

Baron, R. S., Kerr, N. L., and Miller, N. (1992). *Group Process, Group Decision, Group Action*. Buckingham: Open University Press.

Béal, M., and Grégoire, Y. (2021). How do observers react to companies' humorous responses to online public complaints? *J. Serv. Res.* 25, 1–18. doi: 10.1177/1094670521989448

Beard, F. K. (2008). Advertising and audience offense: The role of Intentional humor. *J. Mark. Commun.* 14, 1–17. doi: 10.1080/13527260701467760

Berendt, J., Uhrich, S., and Thompson, S. A. (2018). Marketing, Get ready to rumble—How rivalry promotes distinctiveness for brands and consumers. *J. Bus. Res.* 88, 161–172. doi: 10.1016/j.jbusres.2018.03.015

Berger, J., and Milkman, K. L. (2012). What makes online content viral? *J. Mark. Res.* 49, 192–205. doi: 10.1509/jmr.10.0353

Billig, M. (2005). *Laughter And Teasing: Towards A Social Critique Of Humour*. Thousand Oaks: Sage. doi: 10.4135/9781446211779

Brennan, R., Halliday, D., and Tafesse, W. (2015). Content strategies and audience response on Facebook brand pages. *Mark. Intell. Plan.* 33, 927–943. doi: 10.1108/MIP-07-2014-0135

Brodie, R. J., Hollebeek, L. D., Jurić, B., and Ilić, A. (2011). Customer engagement: Conceptual domain, fundamental propositions, and implications for research. *J. Serv. Res.* 14, 252–271. doi: 10.1177/1094670511411703

Campo, S., Askelson, N. M., Spies, E. L., Boxer, C., Scharp, K. M., and Losch, M. E. (2013). Wow, that was funny' the value of exposure and humor in fostering campaign message sharing. *Soc. Mark. Q.* 19, 84–96. doi: 10.1177/1524500413483456

Cheung, M. L., Pires, G. D., and Rosenberger, I. I. I. P. J. (2019). Developing a conceptual model for examining social media marketing effects on brand awareness and brand image. *Int. J. Econ. Bus. Res.* 17, 243–261. doi: 10.1504/IJEBR.2019.10019330

Cline, T. W., Altsech, M. B., and Kellaris, J. J. (2003). When does humor enhance or inhibit Ad responses?-the moderating role of the need for humor. *J. Advert.* 32, 31–45. doi: 10.1080/00913367.2003.10639134

Coelho, R. L. F., de Oliveira, D. S., and de Almeida, M. I. S. (2016). Does social media matter for post typology? impact of post content on Facebook and Instagram metrics. *Online Inf. Rev.* 40, 458–471. doi: 10.1108/OIR-06-2015-0176

Dessart, L., Veloutsou, C., and Morgan-Thomas, A. (2015). Consumer engagement in online brand communities: A social media perspective. *J. Prod. Brand Manag.* 24, 28–42. doi: 10.1108/JPBM-06-2014-0635

Ducharme, L. J. (1994). Sarcasm and interactional politics. *Symb. Interact.* 17, 51–62. doi: 10.1525/si.1994.17.1.51

Eisend, M. (2011). How humor in advertising works: A meta-analytic test of alternative models. *Mark. Lett.* 22, 115–132. doi: 10.1007/s11002-010-9116-z

Eisend, M. A. (2009). Meta-analysis of humor in advertising. *J. Acad. Mark. Sci.* 37, 191–203. doi: 10.1037/0033-2909.112.1.125

Fornell, C., and Larcker, D. F. (1981). Structural equation models with unobservable variables and measurement error: Algebra and statistics. *J. Mark. Res.* 18, 382–388. doi: 10.1177/002224378101800313

Gervais, M., and David, S. W. (2005). The evolution and functions of laughter and humor-A synthetic approach. *Q. Rev. Biol.* 80, 395–430. doi: 10.1086/498281

Gregory, G. D., Crawford, H. J., Lu, L., and Ngo, L. (2019). Does humour travel? Advertising practices and audience effects in the United States and People's Republic of China. *Int. J. Advert.* 38, 957–978. doi: 10.1080/02650487.2019.1590068

Haugh, M. (2014). Jocular mockery as interactional practice in everyday anglo-Australian conversation. *Aust. J. Linguist.* 34, 76–99. doi: 10.1080/07268602.2014.875456

Holbert, R. L., Hmielowski, J., Jain, P., Lather, J., and Morey, A. (2011). Adding nuance to the study of political humor effects: Experimental research on juvenalian satire versus horatian satire. *Am. Behav. Sci.* 55, 187–211. doi: 10.1177/0002764210392156

Jargon, J. (2017). You want snark with those fries? No one is safe from Wendy's tweets. Wall Street Journal (Online).

Java, A., Song, X., Finin, T., and Tseng, B. (2007). "Why we twitter: understanding Microblogging usage and communities," in *Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 Workshop on Web Mining and Social Network Analysis* (New York, NY), 56–65. doi: 10.1145/1348549.1348556

Kaplan, A. M., and Haenlein, M. (2011). Two hearts in three-quarter time: How to waltz the social media/viral marketing dance. *Bus. Horiz.* 54, 253–263. doi: 10.1016/j.bushor.2011.01.006

Kowalski, R. M. (2004). Proneness to, perceptions of, and responses to teasing: The influence of both intrapersonal and interpersonal factors. *Eur. J. Pers.* 18, 331–349. doi: 10.1002/per.522

Kwak, H., Puzakova, M., and Rocereto, J. F. (2015). Better not smile at the price: The differential role of brand anthropomorphization on perceived price fairness. *J. Mark.* 79, 56–76. doi: 10.1509/jm.13.0410

Li, L. J., and Wang, F. (2022). Disparagement humor: Could laughter dissolve hostility? *Adv. Psychol. Sci.* 30, 670–683. doi: 10.3724/SP.J.1042.2022.00670

MacInnis, D. J., and Folkes, V. S. (2017). Humanizing brands: When brands seem to be like me, part of me, and in a relationship with me. *J. Consum. Psychol.* 27, 355–374. doi: 10.1016/j.jcps.2016.12.003

Martin, A. R., Puhlik-Doris, P., Larsen, G., Gray, J., and Weir, K. (2003). Individual differences in uses of humor and their relation to psychological well-being: Development of the humor styles questionnaire. *J. Res. Pers.* 37, 48–75. doi: 10.1016/S0092-6566(02)00534-2

McGraw, A. P., and Warren, C. (2010). Benign violations: Making immoral behavior funny. *Psychol. Sci.* 21, 1141–1149. doi: 10.1177/0956797610376073

McGraw, A. P., Schwartz, J. A., and Tetlock, P. E. (2012a). From the commercial to the communal: Reframing taboo trade-offs in religious and pharmaceutical marketing. *J. Consum. Res.* 39, 157–173. doi: 10.1086/662070

McGraw, A. P., Warren, C., Williams, L. E., Williams, L., and Leonard, B. (2012b). Too close for comfort, or too far to care? Finding humor in distant tragedies and close mishaps. *Psychol. Sci.* 23, 1215–1223.

McGraw, A. P., Warren, C., and Kan, C. (2015). Humorous complaining. *J. Consum. Res.* 41, 1153–1171. doi: 10.1016/j.jpsychores.2006.12.007

Nair, A. S. (2016). Effect of humour on eWOM intention: The moderating role of product involvement and emotional attachment. *Shanlax Int. J. Manag.* 4, 50–60.

Newton, J. D., Wong, J., and Newton, F. J. (2016). Listerine–for the bridesmaid who's never a bride: Disparaging humour increases brand attitude and recall among the powerless. *Eur. J. Mark.* 50, 1137–1158.

Nielsen. (2015). *Global Trust In Advertising: Winning Strategies For An Evolving Media Landscape*. Available Online at: https://www.nielsen.com/content/dam/nielsenglobal/apac/docs/reports/2015/nielsen-global-trust-in-advertising-report-september-2015.pdf (accessed Oct 3, 2016).

Park, C., and Lee, T. M. (2009). Information direction,Website reputation and eWOM effect: A moderating role of product type. *J. Bus. Res.* 62, 61–67. doi: 10.1016/j.jbusres.2007.11.017

Pentina, I., Guilloux, V., and Micu, A. C. (2018). Exploring social media engagement behaviors in the context of luxury brands. *J. Advert.* 47, 55–69. doi: 10.1080/00913367.2017.1405756

Petter, S., Straub, D., and Rai, A. (2007). Specifying formative constructs in information systems research. *MIS Q.* 31, 623–656. doi: 10.2307/25148814

Pinkleton, B. E., Um, N. H., and Austin, E. W. (2002). An exploration of the effects of negative political advertising on political decision making. *J. Advert.* 31, 13–25. doi: 10.1080/00913367.2002.10673657

Rahman, K. (2013). 'Wow! It's cool': The meaning of coolness in marketing. *Mark. Intell. Plan.* 31, 620–638. doi: 10.1108/MIP-09-2012-0094

Romero, E. J., and Cruthirds, K. W. (2006). The use of humor in the workplace. *Acad. Manag. Perspect.* 20, 58–69. doi: 10.5465/amp.2006.20591005

Shapiro, J. P., Baumeister, R. F., and Kessler, J. W. (1991). A three-component model of children's teasing: Aggression, Humor, and Ambiguity. *J. Soc. Clin. Psychol.* 10, 459–472. doi: 10.1521/jscp.1991.10.4.459

Shifman, L. (2012). An anatomy of a YouTube meme. *New Media Soc.* 14, 187–203. doi: 10.1177/1461444811412160

Smith, P. K., Mahdavi, J., Carvalho, M., Fisher, S., Russell, S., and Tippett, N. (2008). Cyberbullying: Its nature and impact in secondary school pupils. *J. Child Psychol. Psychiatry* 49, 376–385. doi: 10.1111/j.1469-7610.2007.01846.x

Speck, P. S. (1991). The humorous message taxonomy: A framework for the study of humorous ads. *Curr. Issues Res. Advert.* 13, 1–44.

Sundar, A., and Noseworthy, T. J. (2016). Too exciting to fail, too sincere to succeed: The effects of brand personality on sensory disconfirmation. *J. Consum. Res.* 43, 44–67. doi: 10.1093/jcr/ucw003

Svebak, S. (1996). The development of the sense of humor questionnaire:from SHQ to SHQ-6. *Humor* 9, 341–362. doi: 10.1515/humr.1996.9.3-4.341

Swaminathan, V., Stilley, K. M., and Ahluwalia, R. (2009). When brand personality matters: The moderating role of attachment styles. *J. Consum. Res.* 35, 985–1002. doi: 10.1086/593948

Taecharungroj, V., and Nueangjamnong, P. (2014). "The effect of humour on virality: the study of Internet memes on social media," in *7th International Forum on Public Relations and Advertising Media Impacts on Culture and Social Communication* (Bangkok).

Thomas, V. L., and Fowler, K. (2021). Turn the other cheek or an eye for an eye: Exploring brand-to-brand dialogue on social media. *J. Advert.* 50, 441–460. doi: 10.1080/00913367.2020.1867262

Toubia, O., and Stephen, A. T. (2013). Intrinsic versus image-related motivations in social media: Why do people contribute content to twitter? *Mark. Sci.* 32, 368–392. doi: 10.1287/mksc.2013.0773

Warren, C., and McGraw, A. P. (2015). Opinion: What makes things humorous. *Proc. Natl. Acad. Sci. U.S.A.* 112, 7105–7106. doi: 10.1073/pnas.1503836112

Warren, C., Barsky, A., and McGraw, A. P. (2018). Humor, comedy, and consumer behavior. *J. Consum. Res.* 45, 529–552. doi: 10.1093/jcr/ucy015

Warren, C., Carter, E. P., and McGraw, A. P. (2019). Being funny is not enough: The influence of perceived humor and negative emotional reactions on brand attitudes. *Int. J. Advert.* 38, 1025–1045. doi: 10.1080/02650487.2019.1620090

Wright, C. N., and Roloff, M. E. (2013). The influence of type of teasing and outcome on the negative experiences of teasers. *Hum. Commun.* 16, 95–107.

frontiers | Frontiers in Communication

TYPE Brief Research Report
PUBLISHED 18 August 2022
DOI 10.3389/fcomm.2022.910745

Check for updates

OPEN ACCESS

EDITED BY
Donna Mae Erickson,
Haskins Laboratories, United States

REVIEWED BY
Puyang Geng,
Academy of Forensic Science, China
Marisa Cruz,
University of Lisbon, Portugal

*CORRESPONDENCE
Jana Voße
j.vosse@uni-bielefeld.de

SPECIALTY SECTION
This article was submitted to Language Sciences, a section of the journal Frontiers in Communication

RECEIVED 01 April 2022
ACCEPTED 27 July 2022
PUBLISHED 18 August 2022

CITATION
Voße J, Niebuhr O and Wagner P (2022) How to motivate with speech. Findings from acoustic phonetics and pragmatics.
Front. Commun. 7:910745.
doi: 10.3389/fcomm.2022.910745

COPYRIGHT
© 2022 Voße, Niebuhr and Wagner. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

How to motivate with speech. Findings from acoustic phonetics and pragmatics

Jana Voße[1,2]*, Oliver Niebuhr[3] and Petra Wagner[1,2]

[1]Phonetics Work Group, Bielefeld University, Bielefeld, Germany, [2]CITEC, Bielefeld University, Bielefeld, Germany, [3]Centre for Industrial Electronics, University of Southern Denmark, Sonderborg, Denmark

In the present work, we describe and discuss two studies in the field of motivating-speech research. The studies investigate voice-quality features (study 1) and pragmatic aspects (study 2) in German motivating speech, thereby adding to the current state-of-the-art in understanding motivating speech and language[1]. We find indications that a low amount of breathiness, a more periodic signal and a balanced distribution of specific pragmatic elements contribute to a motivating impact in German.

KEYWORDS

motivation, speaking style, acoustic phonetics, pragmatics, voice quality

Introduction

Be it in teaching, in the workplace, in sports or health care, there are numerous scenarios in which it is essential that an individual motivates their interlocutor toward a specific action. In this work, we focus on motivation in the context of sports and healthy nutrition and define the term motivation as a process which results in a readiness to act within an individual (Mook, 1987, cited in Rudolph, 2013). This process is complex and can be influenced by several factors, among which are leadership communication (Mayfield and Mayfield, 2018) and phonetics (Voße and Wagner, 2018). It has been discussed and partially shown that specific patterns on the level of language and speech influence the motivating impact of a person positively, while other patterns are assumed to be detrimental in the context of motivation.

Mayfield and Mayfield (2018) investigate the role of language within motivation from a pragmatic perspective. Central in their work is the Motivating Language Theory by Sullivan (1988) which argues that motivation within the workplace is enhanced by communication from the leadership that uses and balances explaining elements [*meaning making language* (MM)], empathetic elements [*empathetic language* (E)], and instructive elements [*direction-giving language* (DG)]. MM, E, and DG are claimed to be essential and, therefore, should all occur in motivating language.

Voße and Wagner (2018) focus on the acoustic-phonetic expression of motivation. Their hypotheses are inspired by recent studies on emotional speech and charismatic speech. Voße and Wagner (2018) see charisma and motivation as related

1 We acknowledge support for the publication costs by the Open Access Publication Fund of Bielefeld University and the Deutsche Forschungsgemeinschaft (DFG).

concepts, as both can evoke an inner change within an individual (Niebuhr and Gonzalez, 2019). Although both concepts overlap in some aspects, charisma is usually described in the context of leadership (Klein and House, 1995; Antonakis et al., 2016). Motivation is more multifaceted and not necessarily tied to a leader-follower-relation (see Rheinberg and Engeser, 2018 for an overview). Due to the pragmatic proximity of motivation and charisma, Voße and Wagner (2018) expect that acoustic-phonetic characteristics observed for charismatic speech may also apply toward motivating speech.

Regarding motivation and emotion, Voße and Wagner (2018) argue for a systematic connection between these concepts. Striving for a positive emotional state is a basic motive within human behavior and it plays a central role in motivational processes (Schultheiss and Wirth, 2018). The emotional state of an individual, in turn, can directly be influenced by an interlocutor via the display of specific emotions ("emotional contagion", Giner-Sorolla, 2012). Following from this, Voße and Wagner (2018) assume that the expression of positive emotions supports the motivating impact of a speaker.

Consistent with their assumption, Voße and Wagner (2018) found a high and variable pitch, a fast articulation, and a consistently loud voice within the speech units they analyzed as well as a longer duration of the analyzed speech units. This feature set shows several parallels to the acoustic-phonetic expression of charisma and positive emotions (e.g., Burkhardt and Sendlmeier, 2000; Niebuhr et al., 2016).

Understanding the characteristics of motivating speech, however, is still far from being achieved. In the present work, we want to contribute to the current state-of-the-art by presenting our analyses of voice-quality features (study 1) and pragmatic aspects (study 2) in German motivating speech. We designed the two studies with the aim to provide a first insight into these topics, so our findings might constitute a basis for future studies in this area.

Voice quality features have not been investigated in motivating speech so far. We assume these features add to the expression of motivation, as voice quality has been shown to be involved in the acoustic-phonetic characterization of emotions, affect, moods, and attitudes. However, the results regarding positive emotional speech from these studies are not congruent. For example, while Murray and Arnott (1993) report a rather breathy voice for the expression of happiness, Gobl and Ní Chasaide (2003) report a tenser voice setting.

Besides its role in positive emotional speech, Niebuhr et al. (2018) found specific voice quality characteristics for charismatic speech. Following Voße and Wagner's (2018) understanding of charisma and motivation as intertwined concepts, we expect that voice quality characteristics observed for charismatic speech may also apply toward motivating speech.

Study 1 of this work focuses on voice quality features in motivating speech in German. As the picture regarding voice quality characteristics of positive emotional speech is less clear in the literature than the picture regarding voice quality features of charismatic speech, we consider only findings of the latter for the hypothesis to be tested in study 1:

- H1 (study 1): Motivating speech involves a specific voice quality, that is, a bundle of acoustic-phonetic voice quality features all of which conspire to contribute to the impression of a "fuller voice" and a lower amount of breathiness (see Niebuhr et al., 2018, and 2.2 for an explanation of "full").

To test this hypothesis 1, we compare speech samples from two different levels of motivation, namely more motivating speech (MMS) and less motivating speech (LMS) (see Data collection and annotation for a definition of MMS and LMS).

In addition to investigating the role of voice quality in the expression of motivation, it would be insightful to see whether claims of the Motivating Language Theory generalize to domains outside of leadership communication in the workplace. We dedicate study 2 in this article to this purpose. Here, we transfer a central claim of Motivating Language Theory–namely that an occurrence of the element types MM, E, and DG is essential for a motivating impact–to motivating speech in the area of sports and healthy nutrition. To this purpose, we compare the distributions of the element types in more motivating speech (MMS) and in less motivating speech (LMS) and hypothesize the following:

- H2 (study 2): We assume that all three elements MM, E, and DG occur in the MMS condition, but not or less balanced in the LMS condition.

We are further interested in whether the individual element types show specific acoustic-phonetic profiles so that their variable combinations cause dynamic changes within motivating speech. Findings from Repp (2020) show that different types of speech acts are marked differently in the prosodic dimension, which lets us assume that also other pragmatic units such as the element types MM, E, and DG might differ in their acoustic-phonetic realization. Additionally, we expect to observe differences in the realization of each element type between the conditions MMS and LMS, as acoustic-phonetic differences between these two conditions have been reported by Voße and Wagner (2018) for other types of units. In sum, we expect to observe the following regarding the acoustic-phonetic form of the element types MM, E, and DG:

- H3 (study 2): We will observe element-specific, differentiated acoustic-phonetic characteristics between the conditions MMS and LMS for the elements MM, E, and DG.

Methods

Data collection and annotation

We use audio tracks from non-professional motivating videos from the Internet that were freely accessible. We opt for non-professionals to avoid any confounds caused by an expert status of a speaker or due to monetary reasons. Additionally, we ensure that the data is suitable for acoustic analyses and that factors which potentially influence motivational processes are constant. We decide to focus on female motivating speakers of a defined age span (∼18–30 years), who address a specific audience (young adults, approximately 18–30 years, probably mostly females). The assumptions about the target audience are made on the basis of the topics/arguments provided in the videos. Furthermore, we consider only data which show a similar argumentation structure and do not make exhaustive use of visualization tools.

We use the audience rating "view" to distinguish between more motivating videos [representing *more motivating speech* (MMS)] and videos with a less motivating impact [representing *less motivating speech* (LMS)], in this way identifying acoustic-phonetic patterns that potentially either promote or inhibit a motivating impact. In our scheme, more motivating videos gain at least 33,000 views and less motivating videos at most 1,000 views over approximately 1.5 years. These thresholds do not refer to any standard. They are chosen after inspecting the views of all videos suitable for the analyses with the intention to maximize the distance between the MMS and LMS conditions.

We are aware that the popularity of a speaker is a potential confound in our categorization strategy that limits our conclusions. To counteract that, we chose videos of non-professionals who do not receive a high level of recognition outside the fitness community. Doing so, we assume that the majority of the users who watched the analyzed videos did not do so because of the speakers but because of a genuine interest in sports and healthy nutrition. Of course, we cannot fully control the factor popularity with this approach, as it might be that more popular fitness bloggers receive more views than less popular ones. However, we believe that the popularity of a fitness blogger is at least partially founded in its motivating impact. Accordingly, it appears rather unlikely that a video receives "views" exclusively for the popularity of the speakers without considering their motivating impact.

In total, the data set for studies 1 and 2 consists of audio tracks from six videos, each presented by a different female speaker in German. The data is divided into three videos (35 min in total) that represent a higher level of motivation and three videos (18 min in total) that represent a lower level. In total, about 53 min of speech material were analyzed.

The data for the study on voice quality is annotated semi-automatically with WebMAUS (Kisler et al., 2017) on phone level, followed by manual corrections of phone labels and segment boundaries.

For the study on pragmatics, we manually annotated the data for the element types MM, E and DG and supplemented this annotation with an additional, parallel annotation of speech acts from Speech Act Theory (Austin, 1975). The latter facilitated the annotation procedure, as the temporal extension of MM, E and DG elements is often hard to define, whereas speech acts are more clear-cut categories along the time axis and easy to associate with MM, E and DG. From the categories provided by the Speech Act Theory, we considered *representative speech acts*, *direct speech acts* and *indirect speech acts*, as they closely match with the element types MM, E, and DG (see Sullivan, 1988 and Austin, 1975 for more detailed information on the element types and speech act types).

As we applied the Motivating Language Theory outside the field it was originally developed for, it was necessary to adapt the element types MM, E, and DG accordingly. In the present study, MM, E, and DG are defined as follows:

- Meaning making language (MM): representative speech act expressing factual knowledge or personal opinions; Example: "It is fun to do sports."
- Empathetic language (E): representative speech act expressing personal experience; Example: "I really enjoy doing sports."
- Direction giving language (DG): (in)direct speech act expressing commands or suggestions for an action; Example: "Give it a try to do sports."

These definitions form the basic segmentation criteria during the annotation process of study 2.

Data analyses

Study 1: Voice quality

As specified in the introduction, we expect motivating speech to resemble charismatic speech in terms of voice quality. That is, motivating speech should show a low amount of breathiness and a fuller voice (Niebuhr et al., 2018). As acoustic correlates of amount of breathiness and full voice, we consider the features cepstral peak prominence smooth (CPPS), spectral slope, and H1-H2.

CPPS describes the distance between the amplitude of the first rhamonic and its corresponding quefrency point on the regression line in the cepstrum (Hillenbrand and Houde, 1996). This distance is an indicator of the harmonicity of the spectrum, as a prominent first rhamonic speaks for a more clearly pronounced harmonic structure in the spectrum (Mayer, 2017). A more harmonic structure in the spectrum implies a higher periodicity of the signal, for which a lower amount of breathiness is a prerequisite (Hillenbrand and Houde, 1996). We

assume that signals with a high periodicity and a low amount of breathiness are more strongly associated with the auditory impression of a "fuller voice" than signals with a low periodicity and a higher amount of breathiness.

Breathy voices are furthermore characterized by long open phases and short closed phases of the glottis, while the opposite is true for creaky voices (Mayer, 2017). Long open phases of the glottis are usually associated with a larger energy drop in higher frequency bands in comparison to shorter opening phases. The amount of this energy loss toward higher spectral frequencies can be estimated *via* measures of the spectral slope (Mayer, 2017). We use the Praat (Boersma and Weenink, 2018) command "Get slope…" (settings: low band: 50–1,000 Hz, high band: 1,000–5,000 Hz; averaging method: energy) based on a long-term average spectrum.

Besides its relation to breathiness, studies found that the spectral slope also varies due to stress and accentuation, with lower stress/accent levels showing steeper spectral slopes than higher stress/accents levels of syllables and words (see Heldner, 2001 for an overview). Yarra et al. (2017) report a relation between stressed syllables and sonority such that sonority-based features improved the accuracy of their stress-detection algorithm. We infer from these findings that a less steep spectral slope correlates with a more sonorous voice and hence contributes to the impression of a "fuller voice."

With H1-H2, we include a central, robust measure of breathiness in our analysis. H1-H2 describes the dB difference in amplitude between the first and the second harmonic in the signal, which directly relates to the proportion of open and closed phases of the glottis (Mayer, 2017) on the acoustic level.

We expect CPPS, spectral slope, and H1-H2 to differ significantly between the conditions MMS and LMS in the following ways:

- CPPS: higher for MMS than for LMS, that is, a higher signal periodicity, a lower amount of breathiness and, therefore, a fuller voice
- Spectral slope: lower for MMS than for LMS, that is, a lower amount of breathiness and a fuller voice
- H1-H2: closer to 0 for MMS than for LMS, that is, a lower amount of breathiness.

Our analysis is performed exclusively on instances of /aː/ ($n = 246$) with a duration of at least 0.05 s (50 ms), as Mayer (2017) finds this vowel to be most suitable for H1-H2 analyses. To have a consistent data set for this analysis, we took the other measurements on these vowel-sized segments as well.

The analysis is conducted automatically with Praat (Boersma and Weenink, 2018) scripts[2]. The statistical analysis is performed in R. For each of the three voice quality parameters specified in our hypothesis, we determine the following three fundamental statistical features and compare them between LMS and MMS:

- Location of distribution, that is, where the center of the distribution is located
- Dispersion of distribution, that is, the range of the values in a distribution
- Form of distribution, that is, how the distribution is shaped.

During the analysis, we collect visual impressions from the inspection of density plots and supplement them with findings from Bonferroni-corrected non-parametric tests [Wilcoxon-Rank-Sum in the package "stats" (R Core Team., 2019) for location of distribution, Brown-Forsythe in the package "lawstat" (Gastwirth et al., 2015) for dispersion of distribution, Kolmogorov-Smirnov in the package "stats" (R Core Team., 2019) for form of distribution] due to a non-normal distribution of half of the variables.

Study 2: Pragmatics

The data is annotated manually for the element types MM, E, and DG as well as for speech acts (see section Introduction and Data collection and annotation for further information) on a separate tier by one trained phonetician. To check for a consistent annotation procedure, we calculate an intra-annotator agreement and an inter-annotator agreement with Cohen's kappa on a subset of the data (about 10% of the data per speaker). For the intra-annotator agreement, the annotator redoes the labeling task on an already completed subset of the data about 1 year and 3 months later. For the inter-annotator agreement, we delete the labels from a subset of the data annotated by the first annotator, so the annotation contains only boundaries. A second annotator then inserts the respective MM, E and DG element-type and speech-act labels again.

The analysis is conducted in two steps. First, we look at the distribution of elements within the conditions MMS and LMS by applying a Chi-squared test in R (see R Core Team, 2021). Second, we analyze a set of acoustic-phonetic features for each element type automatically via Praat (Boersma and Weenink, 2018) scripts[3]. The set of features comprises:

- Duration of the element types MM, E, and DG (sec.)
- RMS intensity (Pa)
- Intensity variation coefficient ($\frac{\sigma}{\bar{x}}$)
- F0 median (st rel. to 200 Hz)
- F0 variation coefficient ($\frac{\sigma}{\bar{x}}$).

2 We thank Marcin Wlodarczak for sharing his Praat (Boersma and Weenink, 2018) script for the voice quality analysis with us.

3 The scripts were written by the first author, who took inspiration from Mietta Lennes' Praat (Boersma and Weenink, 2018) scripts shared on https://lennes.github.io/spect/.

The statistical assessment is conducted in R on z-score normalized data per speaker. Here, we employ generalized mixed models with the package "lme4" (Bates et al., 2015) to predict the acoustic-phonetic features as a function of Condition*Label. By investigating this interaction effect, we check for element-specific, differentiated acoustic-phonetic characteristics between the conditions for the elements MM, E, and DG as specified in the hypotheses.

Results

Study 1: Voice quality

Regarding *location of distribution*, a clear difference in location of distribution between the conditions MMS and LMS can only be observed for CPPS, where the mean for MMS is noticeably higher than the mean for LMS, see Figure 1. This observation is confirmed by the Wilcoxon rank-sum test, which yields a significant result only for CPPS ($W = 1970$, $p < 0.001^{***}$). On an auditory level, the observed difference in CPPS level corresponds to the impression of a fuller and less breathy voice in MMS than in LMS.

The *dispersion of distribution* differs most between LMS and MMS for H1-H2, where LMS shows a wider dispersion. Differences in this respect are not clearly visible for CPPS and spectral slope, see Figure 1. The Brown-Forsythe test supports this impression, as only the difference in H1-H2 is statistically significant ($F = 41.312$, $p < 0.001^{***}$). This difference indicates that LMS speakers employ a wider spectrum of voice qualities than MMS speakers.

In terms of *form of distribution*, we observe clear differences in the distributions for CPPS and H1-H2, see Figure 1. The statistical significance between these differences is confirmed by the Kolmogorov-Smirnov test ($D = 0.37597$, $p < 0.001^{***}$; $D = 0.36099$, $p < 0.01^{**}$), which supports the impression that LMS speakers behave differently than MMS speakers with respect to these two features.

Study 2: Pragmatics

Regarding the annotator agreement, the calculation of Cohen's kappa yields $\kappa = 0.98$ for the intra-annotator comparison and $\kappa = 0.77$ for the inter-annotator comparison.

For investigating hypothesis 2, we compare the frequencies of the elements MM, E, and DG in the conditions LMS and MMS. Instances of all three elements can be found in both conditions. However, the elements show different distributions in MMS and LMS, χ^2 $(2, N = 442) = 64.386$, $p < 0.001^{***}$. In Figure 2A it can be seen that in MMS the elements are distributed in a more balanced manner compared to LMS where E dominates and DG is hardly present.

Regarding hypothesis 2, we employ linear mixed effect models [package "lme4" (Bates et al., 2015)] to investigate the potential interaction effect Condition*Label. The models contain the variable Speaker as random intercept. An interaction effect is observed in the model for RMS intensity between LMS and E (estimate $= -1.0151$, SE $= 0.4793$, $t = -2.118$). This effect underlies a lower RMS intensity relative to DG labels in the LMS condition, see Figure 2B. A similar trend occurs in the RMS intensity model between the condition LMS and the label MM (meaning making language) (estimate $= -0.8628$, SE $= 0.4933$, t $= -1.749$), but without statistical significance.

Discussion

In study 1, we observe higher mean values in CPPS for MMS than for LMS. This supports our assumption that motivating speech is characterized by a more periodic signal, which we consider as an indicator of a fuller voice and a lower amount of breathiness (see chapter 2.3.1 for more details). The findings regarding H1-H2 indicate reduced and less variable breathiness in motivating speech. In total, these findings support our hypothesis for study 1, that is, a fuller and less breathy voice characterizes MMS as compared to LMS.

The results from study 1 suggest that voice quality-related features play a role in the acoustic-phonetic expression of motivating speech. However, it must be considered that the data selection for this study has been restricted to occurrences of the vowel /aː/ with a minimal duration of 0.05 s. Future studies on other vowel qualities might lead to different observations.

In study 2, we observe different frequency distributions of the elements MM, E, and DG in the conditions MMS and LMS. The nearly equal proportions of the elements in MMS suggest that a balanced occurrence of the three element types promotes a motivating impact, which is consistent with Mayfield and Mayfield (2018). Focusing on E elements and underrepresenting DG elements, however, could have a detrimental effect on motivation. It would be interesting to see in future studies if this presumably detrimental effect can be caused by any possible imbalance in the elements' distributions, or if it is specifically the underrepresentation of DG and/or the overrepresentation of E that reduces a speaker's motivational impact.

Regarding how well the acoustic-phonetic features determine the elements' tone of voice, it is surprising to find an interaction effect only in one of five investigated models, that is, in the RMS intensity model. Considering findings from the literature regarding prosodic realization in speech acts (e.g., Repp, 2020), we expected to find further element-specific acoustic-phonetic characteristics besides the interaction effect of LMS and E in the RMS intensity model. The lack of further

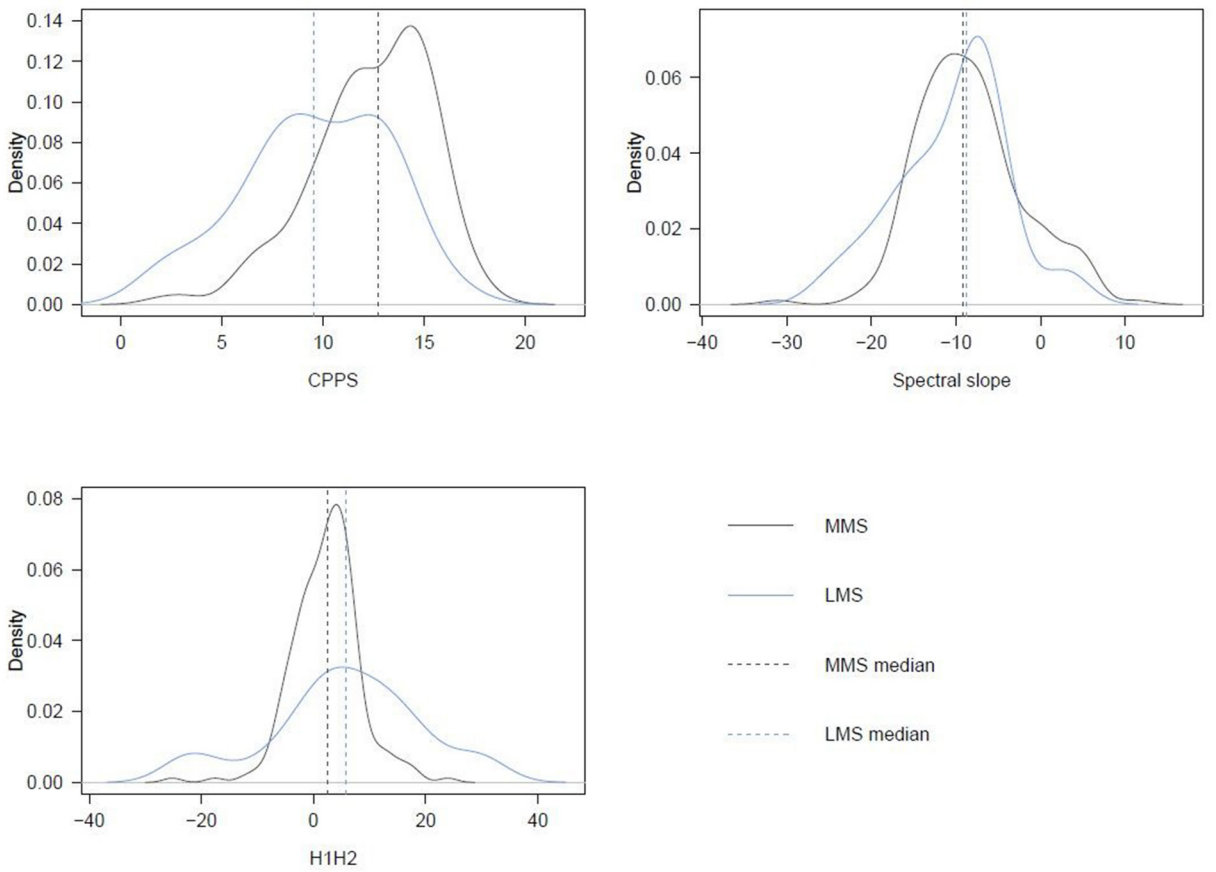


FIGURE 1
Density plots representing the distributions of MMS and LMS for the investigated parameters in study 1.

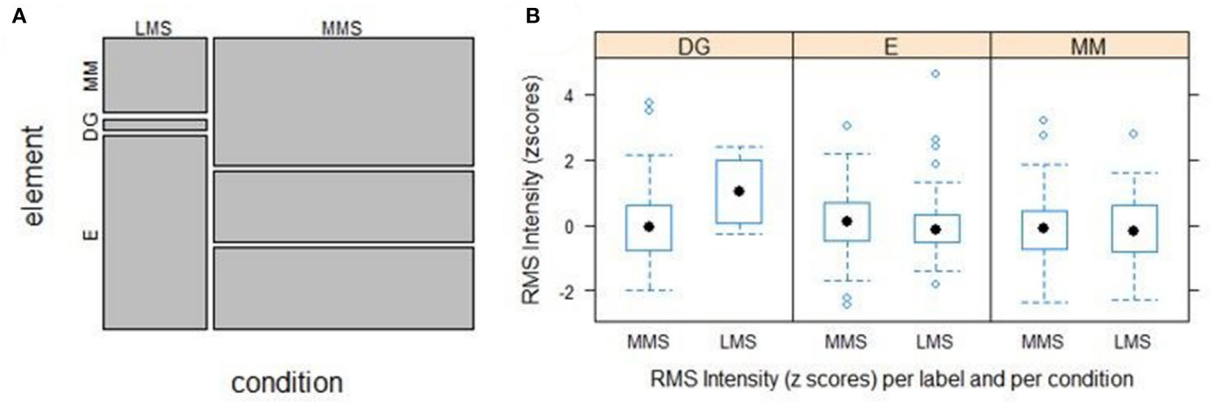


FIGURE 2
(A) Mosaic plot representing the distributions of E, DG, and MM in the conditions MMS and LMS; **(B)** Boxplots for RMS intensity in the elements DG, E and MM with respect to the conditions MMS and LMS.

element-specific observations might be due to our annotation scheme, especially the way we segmented our data. Severe conceptual problems with respect to labeling the data can be ruled out, as the inter-annotator agreement reached a fairly good result ($\kappa = 0.77$). Nevertheless, it would be interesting to run the study again with a revised annotation scheme to

check whether it might lead to a clearer and more distinctive acoustic-phonetic characterization of the three element types. Furthermore, the chosen unit of analysis can have an impact, as Voße and Wagner's (2018) study based on inter-pausal units came up with prosodically more distinct results between the conditions LMS and MMS than our analysis based on the element-type units of study 2.

Another point that must be mentioned is the somewhat arbitrarily chosen number-of-views threshold for grouping the data into the conditions MMS and LMS. We are aware that our threshold lacks a validation and thus limits the generalization of our findings. However, we believe that the chosen threshold captures data with quite different levels of motivation, as we maximized the distance of views between MMS and LMS. We expect this strategy to be sufficient for the present goal: providing a first glimpse into the phonetics and pragmatics of motivating speech and, thereby, laying the foundation for more specific hypotheses that are put forward and tested in future studies. Nevertheless, we believe that follow-up studies in this area will definitely benefit from developing and applying a more robust threshold or categorization concept, for example, one that is based on additional perception tests.

It must be discussed that the data sets of study 1 and 2 are rather sparse and, therefore, the conditions are confounded with a low number of speakers. Also, our samples are restricted to female speakers of a specific age span and, therefore, might not generalize to other groups of speakers. Although we assume that the videos were addressed at young female adults, we have not directly controlled for the audience, so we cannot infer whether this factor might have influenced our results. Moreover, we have not controlled factors such as gestures and mimics, which will certainly have affected the rating of the videos and therefore our threshold concept. These problems result from the general difficulty of obtaining suitable speech material: Not only must potential data sources meet the requirements for acoustic analyses, they must also be constant with respect to factors that play a role in triggering motivation. As only a low number of available data sources fulfill these requirements, future studies might opt for recording larger data sets under controlled conditions.

In view of these limitations, the results of the two studies must be treated as preliminary. Nevertheless, our studies provide a first impression of how motivating speech could possibly be characterized in acoustic-phonetic and pragmatic terms and form a valuable basis for the generation of hypotheses in future studies. In these, besides working with a larger and more controlled data set, we believe that a speaker-specific analysis of motivating speech could be a fruitful endeavor, as there might be differences between individual motivating strategies.

Data availability statement

The data analyzed in this study is subject to the following licenses/restrictions: To protect the privacy of the speakers, we will not publish the links to the material. Requests to access these datasets should be directed to j.vosse@uni-bielefeld.de.

Author contributions

JV: planning, concepting, designing and conducting the studies, and writing and submitting the article. ON: contribution to concept and design of study 2 and feedback on the article draft. PW: supervision of the conduction of the studies, contribution to concept and design of the studies, and feedback on the article draft. JV, ON, and PW: revising the article. All authors contributed to the article and approved the submitted version.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Antonakis, J., Bastardoz, N., Jacquart, P., and Shamir, B. (2016). Charisma: An ill-defined and ill-measured gift. *Annu. Rev. Organ. Psychol. Organ. Behav.* 3, 293–319. doi: 10.1146/annurev-orgpsych-041015-062305

Austin, J. L. (1975). "How to do things with words." In: *The William James lectures delivered at Harvard University in 1955*. Ed. J. H. Urmson and M.Sbisà (New York: Oxford University Press). doi: 10.1093/acprof:oso/9780198245537.001.0001

Bates, D., Maechler, M., Bolker, B., and Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *J. Stat. Softw.* 67, 1–48. doi: 10.18637/jss.v067.i01

Boersma, P., and Weenink, D. (2018). *Praat: doing phonetics by computer [Computer program]. Version 6.1.09.* Available online at: http://www.praat.org/

Burkhardt, F., and Sendlmeier, W. F. (2000). Verification of acoustical correlates of emotional speech using formant-synthesis. In: *SpeechEmotion-2000*, (Newcastle, Northern Ireland, UK), 151–156.

Gastwirth, J. L., Gel, Y. R., Hui, W. W., Miao, W., Noguchi, K., and Lyubchich, M. V. (2015). *Package 'lawstat'*. Available online at: https://cran.microsoft.com/snapshot/2015-07-01/web/packages/lawstat/lawstat.pdf (accessed December 23, 2021).

Giner-Sorolla, R. (2012). *Judging Passions: Moral Emotions in Persons and Groups*. Psychology Press. doi: 10.4324/9780203123874

Gobl, C., and Ní Chasaide, A. (2003). The role of voice quality in communicating emotion, mood and attitude. *Speech Commun.* 40, 189–212. doi: 10.1016/S0167-6393(02)00082-1

Heldner, M. (2001). Spectral emphasis as a perceptual cue to prominence. *TMH-QPSR.* 42, 51–57.

Hillenbrand, J., and Houde, R. A. (1996). Acoustic correlates of breathy vocal quality: Dysphonic voices and continuous speech. *J. Speech Lang. Hear. Res.* 39, 311–321. doi: 10.1044/jshr.3902.311

Kisler, T., Reichel, U. D., and Schiel, F. (2017): Multilingual processing of speech via web services. *Comput. Speech Lang.* 45, 326–347. doi: 10.1016/j.csl.2017.01.005

Klein, K. J., and House, R. J. (1995). On fire: charismatic leadership and levels of analysis. *Leadership Q.* 6, 183–198. doi: 10.1016/1048-9843(95)90034-9

Mayer, J. (2017). *Phonetische Analysen mit Praat. Ein Handbuch für Ein- und Umsteiger. Version 2017/09*.

Mayfield, J., and Mayfield, M. (2018). *Motivating Language Theory: Effective Leader Talk in the Workplace*. New York, NY: Palgrave Macmillan. doi: 10.1007/978-3-319-66930-4

Mook, D. G. (1987). *Motivation: The Organization of Action*. New York, NY: WW Norton.

Murray, I. R., and Arnott, J. L. (1993). Toward the simulation of emotion in synthetic speech: a review of the literature on human vocal emotion. *J. Acoust. Soc. Am.* 93, 1097–1108. doi: 10.1121/1.405558

Niebuhr, O., and Gonzalez, S. (2019). Do sound segments contribute to sounding charismatic? Evidence from a case study of Steve Jobs' and Mark Zuckerberg's vowel spaces. *Int. J. Acoust. Vib.* 24, 343–355. doi: 10.20855/jav.2019.24.21531

Niebuhr, O., Skarnitzl, R., and Tyločkov,á, L. (2018). The acoustic fingerprint of a charismatic voice-Initial evidence from correlations between long-term spectral features and listener ratings. In: *Proc. 18th International Conference of Speech Prosody, Poland,* 359–363. doi: 10.21437/SpeechProsody.2018-73

Niebuhr, O., Voße, J., and Brem, A. (2016). What makes a charismatic speaker? A computer-based acoustic-prosodic analysis of Steve Jobs tone of voice. *Comput. Hum. Behav.* 64, 366–382. doi: 10.1016/j.chb.2016.06.059

R Core Team. (2019). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. Available online at: https://www.R-project.org/

Repp, S. (2020). The prosody of wh-exclamatives and wh-questions in German: Speech act differences, information structure, and sex of speaker. *Lang. Speech.* 63, 306–361. doi: 10.1177/0023830919846147

Rheinberg, F., and Engeser, S. (2018). Intrinsische Motivation und Flow-Erleben. In: *Motivation und Handeln. Springer-Lehrbuch.* Eds J. Heckhausen, H. Heckhausen (Berlin, Heidelberg: Springer). doi: 10.1007/978-3-662-53927-9_14

Rudolph. (2013). *Motivationspsychologie kompakt*. 3rd edition. Belz, Weinheim, Basel.

Schultheiss, O. C., and Wirth, M. M. (2018). Biopsychologische Aspekte der Motivation. In: *Motivation und Handeln. Springer-Lehrbuch.* Eds J. Heckhausen, H. Heckhausen (Berlin, Heidelberg: Springer). doi: 10.1007/978-3-662-53927-9_10

Sullivan, J. (1988). Three roles of language in motivation theory. *Acad. Manage. Rev.* 13, 104–115. doi: 10.5465/amr.1988.4306798

Voße, J., and Wagner, P. (2018). "Investigating the phonetic expression of successful motivation," in *Proceedings of the 9th Tutorial and Research Workshop on Experimental Linguistics*, 117–120. doi: 10.36505/ExLing-2018/09/0028/000361

Yarra, C., Deshmukh, O. D., and Ghosh, P. K. (2017). Automatic detection of syllable stress using sonority-based prominence features for pronunciation evaluation. In: *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP),* 5845–5849. doi: 10.1109/ICASSP.2017.7953277

frontiers | Frontiers in Computer Science

TYPE Perspective
PUBLISHED 26 September 2022
DOI 10.3389/fcomp.2022.885657

Check for updates

OPEN ACCESS

EDITED BY
Oliver Niebuhr,
University of Southern Denmark,
Denmark

REVIEWED BY
Yonghong Yan,
Institute of Acoustics (CAS), China
Simon Stone,
Technical University Dresden,
Germany

*CORRESPONDENCE
Jürgen Trouvain
trouvain@lst.uni-saarland.de

SPECIALTY SECTION
This article was submitted to Human-Media Interaction, a section of the journal Frontiers in Computer Science

RECEIVED 28 February 2022
ACCEPTED 29 August 2022
PUBLISHED 26 September 2022

CITATION
Trouvain J and Weiss B (2022) Thoughts on the usage of audible smiling in speech synthesis applications.
Front. Comput. Sci. 4:885657.
doi: 10.3389/fcomp.2022.885657

COPYRIGHT
© 2022 Trouvain and Weiss. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Thoughts on the usage of audible smiling in speech synthesis applications

Jürgen Trouvain[1]* and Benjamin Weiss[2]

[1]Department of Language Science and Technology, Saarland University, Saarbrücken, Germany, [2]Quality and Usability Lab, Technische Universität Berlin, Berlin, Germany

In this perspective paper we explore the question how audible smiling can be integrated in speech synthesis applications. In human-human communication, smiling can serve various functions, such as signaling politeness or as a marker of trustworthiness and other aspects that raise and maintain the social likeability of a speaker. However, in human-machine communication, audible smiling is nearly unexplored, but could be an advantage in different applications such as dialog systems. The rather limited knowledge of the details of audible smiling and their exploitation for speech synthesis applications is a great challenge. This is also true for modeling smiling in spoken dialogs and testing it with users. Thus, this paper argues to fill the research gaps in identifying factors that constitute and affect audible smiling in order to incorporate it in speech synthesis applications. The major claim is to focus on the dynamics of audible smiling on various levels.

KEYWORDS

speech synthesis, social signaling, computational paralinguistics, smiling, trustworthiness

1. Introduction

Would users appreciate audible smiling in speech produced by machines, be it text-to-speech (TTS) systems on computers, virtual agents or social robots? If yes, how should the synthesis of audible smiling be approached? There is evidence that smiling in human-human interaction is not only visible but also audible (Tartter, 1980; Tartter and Braun, 1994). Visual signals are often seen as primary in face-to-face situations and acoustic signals as secondary. However, in situations restricted to the auditory-acoustic channel, for instance on the telephone, smiling can only be perceived by vocal and consequently acoustic features.

There is a multitude of functions smiling can have in different everyday social settings. Smiling can be interpreted as a marker of friendliness and politeness, it can be used to express amusement and exhilaration, and it is often applied to build trust between speakers because it strongly increases their social likeability. Obviously, there is a great potential for this effective and attractive social signal to be exploited in human-computer interaction (HCI) and we see a need to clarify the complexity of this topic before we start with the development of technical solutions and the testing of their usability. Thus, the aim of this perspective paper is to provide an ordered collection of thoughts on audible smiling in synthetic speech. We identify findings, but also problems, on various levels and

suggest approaches for solutions. These critical thoughts are neither meant as "wishful thinking" nor as a feasibility study, but they are supposed to shed light on a potential asset to technical systems in terms of an increased naturalness in the interaction with a human user. Our thoughts are always directed to the *auditory* channel (without ignoring the dominant visual channel).

For the domain of visual smiles, it is understood that dynamic characteristics of facial and head movements inherently contribute to the production and perception of the social signal categorization that is manifested as visible smile (Frank et al., 2003). Quite relevant for audible smiles, this includes duration, as, e.g., videos of smiling faces rated as amused were longer than those rated as polite (Ambadar et al., 2009). Examining the morphological and dynamic characteristics of smiles is required in order to discriminate various functions of smiles (Rychlowska et al., 2017). For the acoustic channel, however, this change of perspective has not yet taken place. So, our claim is to concentrate on the dynamics of smiled speech. This change needs to happen on various levels, as the challenges are

- *Temporal dynamics:* Since it is highly unlikely that entire utterances are articulated smiling, the factors affecting choice of sections for expressing a social signal with smiling needs to be identified. In a conversation, this comprises when spoken feedback ("u-hu", "m-hm", "yeah") is audibly smiled.
- *Intensity dynamics:* Are there, like with other affective display in speech, degrees and nuances of audible smiling, and what are their regulatory factors? From a perceptual point of view, graded intensity in smiled speech synthesis can be perceived by humans (El Haddad et al., 2015).
- *Social signaling:* Smiling can be analyzed as a referring expression. But it is still unclear, if referents can be distinguished in speech, what they are and which social and affective function is linked with this smile.

To meet these challenges, we propose to deepen the research on the acoustic properties, perception, and interpretation of smiled speech and to identify factors affecting its dynamics, e.g., content, discourse markers, and social function/meaning.

2. Functions of smiling

2.1. Affective-social components

Prototypical associations of smiling are positive affective states such as *happiness* and *joy*, a *good mood* or *contentment*. Smiling can also be used for *seduction* or as an expression of *amusement* (Schröder et al., 1998), but also to mark *irony*. On the recipient's side, smiling belongs to those social signals that can generate the impression of *interest*, *friendliness*, and *intimacy* (Floyd and Erbert, 2003; Krumhuber et al., 2007; Burgoon et al., 2018).

These positively associated types of smiling bear an authentic character. In addition there are non-authentic or not genuinely felt types of smiling. Happiness is probably often expressed with a smile, but a smile is not necessarily linked to positive emotions. Examples include situations where negative emotions are masked with an expression of joy or situations in which individuals have feelings of *uncertainty*, *nervousness* or *embarrassment* (Keltner, 1995). This wide range of meanings of smiling have even been replicated in artificial faces (Ochs et al., 2017). Further examples include the expression of *dominance* toward others. The difference between "authentic/felt" and "non-authentic/non-felt" smiling in the visual channel is mainly reflected by the contraction or non-contraction of the eye-ring muscle (m. orbicularis oculi), the so-called Duchenne-smile (Ekman and Friesen, 1982). This difference in muscle contraction is also the base for separating *trustworthy* from *deceptive* behavior (Ekman and Friesen, 1982). For the visual channel, this fundamental distinction seems to be established, however not for the acoustic channel.

The general impression that a smiling face is regarded as more *attractive* than a non-smiling face is evidenced by numerous studies. For instance, in a Brazilian study smiling faces were considered as happier and even as more attractive than a neutral expression (Otta et al., 1982). Regarding the concept of visual attractiveness, a smile enforces the positive assessment of faces, particularly of females (Lau, 1982). Moreover, females typically smile more when flirting (Moore, 1985). For male faces this effect is not that clear (Mehu et al., 2008; Okubo et al., 2015). The transfer of these findings to synthetic voices is difficult because non-verbal material (photographs and videos), particularly in field research, has no reported relation to co-verbal smiling.

There are several forms of social smiling. They are core features for the display of *politeness* and *friendliness*, but they are not necessarily expressed with a Duchenne marker. A smile can also be used to show *empathy* and *agreement* with somebody else. Studies demonstrate that smiles that were rated as more genuine strongly predict judgments about the *trustworthiness* (Centorrino et al., 2015). In the majority of cases, genuine smiles trigger a reciprocal social action, even in HCI (Krämer et al., 2013).

It would generally be helpful if synthesized speech applications could express social functions associated with a smiling voice when appropriate. However, while state-of-the-art findings (Section 3) and data-driven models (Section 5.1) have reached a level to produce audible smiling, a solid basis to confirm or reject differences in audible smiles that refer to different social functions, is not known to us. In addition, other factors regulating the dynamics and location of audible smiling are yet to be modeled. As an example can serve the reciprocity in terms of initiating smiles and smiling back (Arias et al., 2018).

While the relevance of reciprocate smiling is evident, is still unknown which conversational sections have to be synthesized and how the exact timing of this reciprocal mechanism works. On a broader lever, however, i.e., by treating this mechanism as synchrony, it could be shown that it is observable throughout whole conversations (Rauzy et al., 2022).

This goal is in line with the aims of social signal processing, e.g., formulated by Vinciarelli et al. (2012): "a human-centered vision of computing where intelligent machines seamlessly integrate and support human-human interactions, embody natural modes of human communication for interacting with their users [...] At its heart, social intelligence aims at correct perception, accurate interpretation, and appropriate display of social signals". In our view, this demand is still completely blank with respect to appropriate and perceivable synthesis of smiled speech.

2.2. Cultural interpretations of smiling

It is tempting to assume that positive smiles are always realized as an "authentic" smile with a Duchenne marker. Likewise, it could be assumed that a smile of an unacquainted person is generally perceived as attractive, friendly and definitively positive (e.g., pictures in application letters or on personal homepages). However, there is evidence that in some non-Western cultures an authentic smile is not bound to a Duchenne marker (Thibault et al., 2012). Moreover, in a cross-cultural comparative study investigating face perception with subjects from more than 40 cultures, it could be shown that in some cultures smiling faces of unacquainted persons leave a negative impression on observers (Krys et al., 2016). Thus, smiling *per se* does not necessarily lead to a more positive impression of the perceiver. This could also be the case for audibly transmitted smiling, particularly when coming from a synthesized voice. Another example of cultural diversity regarding the usage of smiling is provided by a study where Chinese and Dutch kindergarten children were asked to play a game—either alone or together with peers (Mui et al., 2017). In contrast to the Dutch children, who did not change their smiling behavior between both conditions, the Chinese children smiled more when playing with other children.

3. Acoustic characteristics of smiled voices

A clear distinction should be made between smiling and laughter. Both concepts can have similar functions and sometimes they are used as synonyms (e.g., the expression "s/he laughed with me" when actually the *smile* of a person was directed to another person). Laughter can occur with much variability and complexity (Truong et al., 2019). Most forms of laughter do not overlap with speech, in contrast to "speech-laughs" (Nwokah et al., 1999; Trouvain, 2001) where laughter occurs while articulating. This "laughed speech" is mainly characterized by a high degree of breathiness together with a vibrato-like voice quality (often only for two syllables) and thus differs from "smiled speech" (Trouvain, 2001; Erickson et al., 2009).

Various studies were able to show that smiling is also perceivable from speech and without visual information (Tartter, 1980; Tartter and Braun, 1994). Utterances produced with a non-emotional mechanical lip spreading are perceived as being more "smiled" than utterances without lip spreading (Robson and Beck, 1999). Perceivable smiled speech can be explained with changes of various acoustic parameters: Compared to non-smiled speech the fundamental frequency (F0) is higher due to a higher overall muscular tension, the second formant (F2) is higher due to a shortened vocal tract from lip spreading and a raised larynx. By articulatory synthesis, it could be verified that all three factors have a perceptual effect, but combined, the audible smiling is stronger (Stone et al., 2022). These effects can also be observed for the high unrounded front vowel [i], in contrast to vowels that are (more) rounded and/or lower and/or further back, e.g., [o]. This reflects Ohala's "i-face" for smiling and "o-face" for threatening (Ohala, 1980, 1984). The described tendencies have been confirmed by later studies (Schröder et al., 1998; Drahota et al., 2008).

The perception of smiling from voice also depends on the perceived intensity. A cross-lingual study (Emond et al., 2016) showed that listeners need more time to recognize a mild smile compared to more intense smiles. The same study also revealed a linguistic advantage for the recognition of audible smiles. Listeners were slower in smiling detection and recognized fewer smiles when they did not share the same accent or the same language as the speakers.

Further studies are needed to achieve a more differentiated overview of phonetic parameters such as intensity, duration and voice quality in smiled speech. A particular focus should be on perception, especially the timing of smiling in dialogues, the perceived intensity of audible smiling, and the cross-modal aspects of smiling perception.

4. Possible applications

Often the motivation of researchers and developers is to make machines more human-like. An example for this positive transfer is a study with a human-like virtual agent where adult subjects smiled longer at the robots when the robots showed some (visual) smile as well. This means that the reciprocal smile was increased on the side of the users (Krämer et al., 2013). However, it is not clear whether human users really benefit from a smiling interaction with a machine. For instance, a study where children (9 years) used social robots as learning tutors showed

that the children achieved better results when the robots did not act in a friendly way (Kennedy et al., 2017). This illustrates the need for shifting from mere synthesis of smiling toward proper manifestation of audible smiles as expression of a specific social meaning, for which a solid basis on constituting factors is required.

4.1. Audiobooks

Audiobooks are a wide field of applications for synthetic speech. In audiobook productions using human voices, direct speech of various characters in fictional literature can either be displayed by different professional speakers or by the same speaker who uses different voice qualities for the characters. In synthesized audiobooks, a given character or situation could be displayed by a "smiled voice". A requirement for an appropriate application of smiled speech synthesis would be a text analysis tool that finds those portions of direct speech where smiling fits. This could either be done by finding words of the semantic field of smiling (e.g., grin, mischievous, friendly) or by a sentiment analysis directed to friendliness, politeness and further functions of smiling.

4.2. Social robots

In contrast to virtual (embodied conversational) agents where a high-quality animation of the facial expression is possible (Ochs et al., 2010), many social robots without a display head like Pepper or Nao, do not have the possibility to generate a visible smile. An audible smile could be helpful as a social signal to avoid an uncanny valley effect. Virtual agents with visual smiling were regarded as friendlier and more attractive than those without, and smiling also enforces the impression of extroversion (Cafaro et al., 2012).

A special dimension is opened up when social robots have children as users, for instance care takers in nursing homes or training dolls for autistic children. In the interaction between children and social robots an increased degree of familiarity and trust seems to be substantial. Important components to achieve this are non-verbal behavior, feedback control and other forms of interaction management (Belpaeme et al., 2018). Smiling, including audible smiling, can also play a relevant role in this context.

4.3. Dialog systems

The coordinated interaction of conversations depends on proper timing of production or even missing production of spoken signals to convey meaning and to ensure the conversational flow. This kind of coordination comprises back-channel and turn-taking signals (Enfield, 2017) as well as (automatic) convergence (Branigan et al., 2010). Both kinds are potentially subject to audible smiles, but only for the latter, empirical evidence is known to the best of our knowledge.

This kind of convergence could also be observed by Krämer et al. (2013) where adult subjects smiled for a longer time with artificial agents when also the agent showed a (visual) smile. This is in line with Torre et al. (2020) who directly tested audible smiling in a gaming scenario and found increased trustworthiness even in contradicting behavioral evidence.

5. Smiled synthetic speech

The challenge for integrating audible smiling in speech synthesis can be regarded at different levels. For different methods of *signal generation* the limited knowledge about audible smiling in humans should be exploited. *Modeling* audible smiling in dialogs requires the control of temporal, discourse-relevant and cultural aspects, in addition to the signal generation. Last but not least the *evaluation* of appropriateness in given applications represents the third component.

5.1. Signal generation

Articulatory synthesis would be an obvious choice for verifying the perceptual validity of findings concerning the properties of smiled speech. Testing and verifying such analytic results for perception, in our domain social signals manifested by smiling and its interplay with phonetic dynamics, testing and verifying such analytic results for perception could greatly benefit from articulatory synthesis to produce intensity and dynamic nuances in a controlled way. However, given the current increase in (and maybe demand for) high signal quality, a data-driven approach seems also advisable. An early attempt of smiled speech synthesis for HMM synthesis, utilizing parallel corpora, confirms the perceptual effect of smiling intensity, but also reveals issues with naturalness (El Haddad et al., 2015). These, however, seem to be overcome in more recent work (Kirkland et al., 2021). Still, the typical limitation of data-driven synthesis, i.e., difficulty to draw conceptual conclusions like identifying relevant factors, is not overcome.

5.2. Modeling

Using smiled synthetic speech in real-world applications requires a contextually appropriate control of the synthesis that considers content and culture when selecting sections to be produced as smiled. An automatic symbolic annotation of those

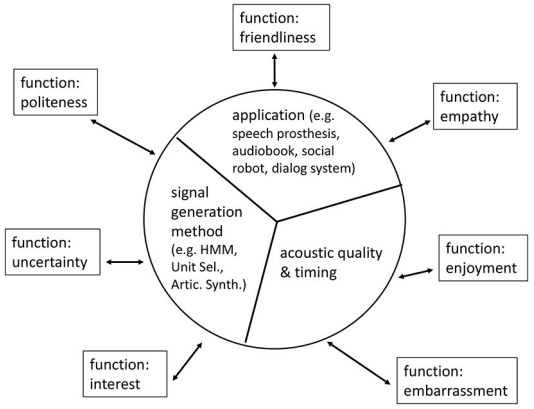


FIGURE 1
Different functions of smiling as evaluation parameters to test the appropriateness of signal generation, application, acoustic quality, and timing.

sections requires a language- and culture-specific model and a sentiment analysis of the text to be synthesized.

Interactive applications could be enhanced by the generation of discourse-dependent social signals, for instance reciprocal smiling. This in return requires a clarification how smiling in human-human interaction is distinct in audio-visual from audio-only situations. In general, we have to model the audible smiling with regard to its timing (duration, start and end relative to speech) and acoustic quality (distinctiveness to non-smiled speech).

5.3. Evaluation

Since smiling has so many functions, an evaluation must monitor the matching of the intention of the generated smile and the interpretation of the perceived smile in given situations. A general assessment with mean opinion scores seems not to be the right way for evaluation but *contextual appropriateness* as demanded by Wagner et al. (2019). Thus, the needs and preferences of the users of synthetic speech must be tested, preferably in a behavioral paradigm, i.e., not (solely) by explicit ratings, but observed behavioral (gamified) choices instead. Smiled voice should not be regarded as one style of expressiveness, but as a carrier mechanism to transport many different expressions—with each expressive function evaluated separately as illustrated in Figure 1.

6. Discussion

There is no doubt that smiling serves a multitude of important social functions. Currently, politeness and friendliness are not yet the core features of synthetic voices. Although the call for more and better expressiveness in speech synthesis has been around for a while (Schröder, 2009), there have been hardly any attempts to tackle this challenge for smiled voices.

For audio books, it can clearly be beneficial to use smiled voices. In dialogical applications, there may be an advantage when users are made aware of smiled voice. Irritations for users evoked by such a human-like unfamiliarity and unexpected peculiarity should be avoided.

Should speaking machines be able to smile? Anthropomorphizing of non-human objects is a possibility which can be applied to building trust. However, it could also lead to disappointed expectations regarding social competence or even to an "uncanny valley of mind" (Gray and Wegner, 2012).

What should be the next steps in the upcoming years (or decades) in order to achieve a more human-like smiling behavior in speech synthesis? There is a research gap for smiled voice in many respects. Research in voice attractiveness still lacks vocal aspects of smiling as an effective mean of sexy, likable and charismatic speakers (Weiss et al., 2020). The majority of research in human smiling is exclusively concentrated on the visual channel. In this research direction, the main objects of study are pictures of faces (often without glasses and beards, and face masks). It lacks the temporal dynamics, the changing intensity of the smile, and the situational and verbal context in which the smile occurred. These can be very important features when modeling smiling in speech (for synthesis or other applications).

In addition, the relation between the visual and the acoustic information is under-explored, particularly in talk-in-interaction (between humans and in HCI). Moreover, situations with machine-aided communication, such as human-robot interaction or as a training device for autistic children, require a thorough understanding of the effects of smiling in the audio-only and the audio-visual modalities as well as in "smile-in-interaction".

Although our thoughts aim at the audio-only aspects of synthesized speech, it can of course also be useful when thinking about audio-visual aspects in speech synthesis, as e.g., in embodied conversational agents or in social robots.

Based on the presented, albeit limited, state description, we argue to fill the identified gaps. While in principle, audible smiling can already be synthesizes for a given duration of speech, the challenges are in quantitative models that incorporate the communicative factors of smiling function (Figure 1), and timing and dynamics of audible smiling in their interrelation to the linguistic and coordinating properties of speech, like smiling duration and intensity within phrases and turns. With such models, we expect a major advance in communicative meaningful synthesis of, e.g., audio books' or artificial agents'

speech. In short, basic research is needed with respect to (i) when exactly, (ii) to which degree, and (iii) for which purpose humans smile in spoken interaction. The research gap also concerns the phonetic aspects of smiled speech. How do acoustics and perception interact? How is visual information processed in combination with acoustic information in speech? How can manipulations be evaluated? Research and development both require more (annotated) data which currently do not exist in the quality and amount needed. We did not regard other social factors, like gender (Hiersch et al., 2022) or status, which are known to affect overall amounts of smiling display, but which we expects not as impactful at this particular state of research. Taken together, we consider our thoughts on audible smiled speech as a contribution that helps to further develop *social signal processing* (Vinciarelli et al., 2012).

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s.

Author contributions

JT and BW: concept, literature research, and formulations. Both authors contributed to the article and approved the submitted version.

Funding

We acknowledge support by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) and Saarland University within the 'Open Access Publication Funding' programme.

Acknowledgments

We thank commentators of a conference presentation given to a similar topic (Trouvain and Weiss, 2020). We are particularly grateful to Iona Gessinger and Mikey Elmers for comments on an earlier draft of this paper.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Ambadar, Z., Cohn, J. F., and Reed, L. I. (2009). All smiles are not created equal: Morphology and timing of smiles perceived as amused, polite, and embarrassed/nervous. *J. Nonverbal Behav.* 1, 17–34. doi: 10.1007/s10919-008-0059-5

Arias, P., Belin, P., and Aucouturier, J.-J. (2018). Auditory smiles trigger unconscious facial imitation. *Curr. Biol.* 28, R782. doi: 10.1016/j.cub.2018.05.084

Belpaeme, T., Vogt, P., van den Berghe, R., Bergmann, K., Goksun, T., Haas M, et al. (2018). Guidelines for designing social robots as second language tutors. *Int. J. Soc. Robot.* 10, 325–341. doi: 10.1007/s12369-018-0467-6

Branigan, H., Pickering, M., Pearson, J., and Mclean, J. (2010). Linguistic alignment between people and computers. *J. Pragmat.* 42, 2355–2368. doi: 10.1016/j.pragma.2009.12.012

Burgoon, J., Buller, D., Hale, J., and Turck, M. (2018). Relational messages associated with nonverbal behaviors. *Hum. Commun. Res.* 10, 351–378. doi: 10.1111/j.1468-2958.1984.tb00023.x

Cafaro, A., Vilhjalmsson, H., Bickmore, T., Heylen, D., Johannsdottir, K., and Valgarosson, G. (2012). "First impressions: users' judgments of virtual agents' personality and interpersonal attitude in first encounters," in *Proc. 12th Int'l Conf. Intell. Virtual Agents* (Santa Cruz, CA), 1–14. doi: 10.1007/978-3-642-33197-8_7

Centorrino, S., Djemai, E., Hopfensitz, A., Milinski, M., and Seabright, P. (2015). Honest signaling in trust interactions: smiles rated as genuine induce trust and signal higher earning opportunities. *Evol. Hum. Behav.* 36, 8–16. doi: 10.1016/j.evolhumbehav.2014.08.001

Drahota, A., Costall, A., and Reddy, V. (2008). The vocal communication of different kinds of smile. *Speech Commun.* 50, 278–287. doi: 10.1016/j.specom.2007.10.001

Ekman, P., and Friesen, W. (1982). Felt, false, and miserable smiles. *J. Nonverbal Behav.* 6, 238–258. doi: 10.1007/BF00987191

El Haddad, K., Cakmak, H., Moinet, A., Dupont, S., and Dutoit, T. (2015). "An HMM-approach for synthesizing amused speech with a controllable intensity of smile," in *IEEE International Symposium on Signal Processing and Information Technology* (Abu Dhabi), 7–11. doi: 10.1109/ISSPIT.2015.7394422

Emond, C., Rilliard, A., and Trouvain, J. (2016). "Perception of smiling in speech in different modalities by native vs. non-native speakers," in *Proc. Speech Prosody*, (Boston, MA), 639–643. doi: 10.21437/SpeechProsody.2016-131

Enfield, N. J. (2017). *How We Talk. The Inner Workings of Conversation*. New York, NY: Basic Books.

Erickson, D., Menezes, C., and Sakakibara, K. (2009). "Are you laughing, smiling or crying?" in *Proc. APSIPA Summit and Conference* (Sapporo), 529–537.

Floyd, K., and Erbert, L. (2003). Relational message interpretations of nonverbal matching behavior: an application of the social meaning model. *J. Soc. Psychol.* 143, 581–597. doi: 10.1080/00224540309598465

Frank, M., Ekman, P., and Friesen, W. (2003). Behavioral markers and recognizability of the smile of enjoyment. *J. Pers. Soc. Psychol.* 64:83–93. doi: 10.1037/0022-3514.64.1.83

Gray, K., and Wegner, D. (2012). Feeling robots and human zombies: mind perception and the uncanny valley. *Cognition* 125, 125–130. doi: 10.1016/j.cognition.2012.06.007

Hiersch, P., McKeown, G., Latu, I., and Rychlowska, M. (2022). "Gender differences, smiling, and economic negotiation outcomes," in *Proceedings of the Workshop on Smiling and Laughter across Contexts and the Life-Span within the 13th Language Resources and Evaluation Conference* (Marseille: European Language Resources Association), 11–15.

Keltner, D. (1995). Signs of appeasement: evidence for the distinct displays of embarrassment, amusement, and shame. *J. Pers. Soc. Psychol.* 68, 441–454. doi: 10.1037/0022-3514.68.3.441

Kennedy, J., Baxter, P., and Belpaeme, T. (2017). The impact of robot tutor nonverbal social behavior on child learning. *Front. ICT Hum. Media Interact.* 4, 6. doi: 10.3389/fict.2017.00006

Kirkland, A., Wå,odarczak, M., Gustafson, J., and Szekely, E. (2021). "Perception of smiling voice in spontaneous speech synthesis," in *Proc. 11th ISCA Speech Synthesis Workshop (SSW 11)*, 108–112. doi: 10.21437/SSW.2021-19

Krämer, N., Kopp, S., Becker-Asano, C., and Sommer, N. (2013). Smile and the world will smile with you – the effects of a virtual agent's smile on users' evaluation and behavior. *Int. J. Hum. Comput. Stud.* 71, 335–349. doi: 10.1016/j.ijhcs.2012.09.006

Krumhuber, E., Manstead, A., Cosker, D., Marshall, D., Rosin, P., and Kappas, A. (2007). Facial dynamics as indicators of trustworthiness and cooperative behavior. *Emotion* 7, 730–735. doi: 10.1037/1528-3542.7.4.730

Krys, K., Vauclair, C.-M., Capaldi, C. A., Miu-Chi Lun, V., Bond, M. H., Domínguez-Espinosa, A., et al. (2016). Be careful where you smile: culture shapes judgments of intelligence and honesty of smiling individuals. *J. Nonverbal Behav.* 40, 101–116. doi: 10.1007/s10919-015-0226-4

Lau, S. (1982). The effect of smiling on person perception. *J. Soc. Psychol.* 117, 63–67. doi: 10.1080/00224545.1982.9713408

Mehu, M., Little, A., and Dunbar, R. (2008). Sex differences in the effect of smiling on social judgments: an evolutionary approach. *J. Soc. Evol. Cult. Psychol.* 2, 103–121. doi: 10.1037/h0099351

Moore, M. (1985). Non-verbal courtship patterns in women: context and consequences. *Ethol. Sociobiol.* 6, 237–247. doi: 10.1016/0162-3095(85)90016-0

Mui, P., Goudbeek, M., Swerts, M., and Hovasapian, A. (2017). Children's non-verbal displays of winning and losing: effects of social and cultural contexts on smiles. *J. Nonverbal Behav.* 41, 67–82. doi: 10.1007/s10919-016-0241-0

Nwokah, E., Hsu, H.-C., Davies, P., and Fogel, A. (1999). The integration of laughter and speech in vocal communication: a dynamic systems perspective. *J. Speech Lang. Hear. Res.* 42, 880–894. doi: 10.1044/jslhr.4204.880

Ochs, M., Niewiadomski, R., and Pelachaud, C. (2010). "How a virtual agent should smile? Morphological and dynamic characteristics of virtual agent's smiles," in *Proc. Int'l Conf. on Intelligent Virtual Agents*, 427–440. doi: 10.1007/978-3-642-15892-6_47

Ochs, M., Pelachaud, C., and McKeown, G. (2017). A user-perception based approach to create smiling embodied conversational agents. *ACM Trans. Interact. Intell. Syst.* 7, 1–33. doi: 10.1145/2925993

Ohala, J. (1980). The acoustic origin of the smile. *J. Acoust. Soc. Am.* 68, S33. doi: 10.1121/1.2004679

Ohala, J. (1984). An ethological perspective on common cross-language utilization of f0 of voice. *Phonetica* 41, 1–16. doi: 10.1159/000261706

Okubo, M., Ishikawa, K., Kobayashi, A., Laeng, B., and Tommasi, L. (2015). Cool guys and warm husbands: the effect of smiling on male facial attractiveness for short- and long-term relationships. *Evol. Psychol.* 13, 1–8. doi: 10.1177/1474704915600567

Otta, E., Abrosio, F. F. E., and Hoshino, R. L. (1982). Reading a smiling face: messages conveyed by various forms of smiling. *Percept. Motor Skills* 82, 1111–1121. doi: 10.2466/pms.1996.82.3c.1111

Rauzy, S., Amoyal, M., and Priego-Valverde, B. (2022). "A measure of the smiling synchrony in the conversational face-to-face interaction corpus PACO-CHEESE," in *Proceedings of the Workshop on Smiling and Laughter across Contexts and the Life-Span within the 13th Language Resources and Evaluation Conference* (Marseille: European Language Resources Association), 16–20.

Robson, J., and Beck, J. M. (1999). "Hearing smiles–perceptual, acoustic and production aspects of labial spreading," in *Proc. 14th Int'l Congress of Phonetic Sciences (ICPhS)* (San Francisco, CA), 219–222.

Rychlowska, M., Jack, R. E., Garrod, O. G. B., Schyns, P. G., Martin, J. D., and Niedenthal, P. M. (2017). Functional smiles: tools for love, sympathy, and war. *Psychol. Sci.* 28, 1259–1270. doi: 10.1177/0956797617706082

Schröder, M. (2009). "Expressive speech synthesis: past, present, and possible futures," in *Affective Information Processing*, eds J. Tao and T. Tan (London: Springer), 111–126. doi: 10.1007/978-1-84800-306-4_7

Schröder, M., Auberge, V., and Cathiard, M.-A. (1998). "Can we hear smiles?" in *Proc. Conference on Spoken Language Processing (ICSLP)* (Sydney, NSW), 559–562. doi: 10.21437/ICSLP.1998-106

Stone, S., Abdul-Hak, P., and Birkholz, P. (2022). "Perceptual cues for smiled voice - an articulatory synthesis study," in *Studientexte zur Sprachkommunikation: Elektronische Sprachsignalverarbeitung 2022*, eds O. Niebuhr, M. S. Lundmark, and H. Weston (Dresden: TUD Press), 131–138.

Tartter, V., and Braun, D. (1994). Hearing smiles and frowns in normal and whisper registers. *J. Acoust. Soc. Am.* 96, 2101–2107. doi: 10.1121/1.410151

Tartter, V. C. (1980). Happy talk: perceptual and acoustic effects of smiling on speech. *Percept. Psychophys.* 27, 24–27. doi: 10.3758/BF03199901

Thibault, P., Levesque, M., Gosselin, P., and Hess, U. (2012). The Duchenne marker is not a universal signal of smile authenticty–but it can be learned! *Soc. Psychol.* 43, 215–221. doi: 10.1027/1864-9335/a000122

Torre, I., Goslin, J., and White, L. (2020). If your device could smile: people trust happy-sounding artificial agents more. *Comput. Hum. Behav.* 105, 106216. doi: 10.1016/j.chb.2019.106215

Trouvain, J. (2001). "Phonetic aspects of 'speech-laughs'," in *Proc. Conference on Orality Gestuality (ORAGE)* (Aix-en-Provence), 634–639.

Trouvain, J., and Weiss, B. (2020). "Uberlegungen zu wahrnehmbarem Lacheln in synthetischen Stimmen," in *31th Conference Elektronische Sprachsignalverarbeitung* (Magdeburg), 26–33.

Truong, K., Trouvain, J., and Jansen, M.-P. (2019). "Towards an annotation scheme for complex laughter in speech corpora," in *Proc. Interspeech* (Graz), 529–533. doi: 10.21437/Interspeech.2019-1557

Vinciarelli, A., Pantic, M., Heylen, D., Pelachaud, C., Poggi, I., D'Errico, F., et al. (2012). Bridging the gap between social animal and unsocial machine: a survey of social signal processing. *IEEE Trans. Affect. Comput.* 3, 69–87. doi: 10.1109/T-AFFC.2011.27

Wagner, P., Beskow, J., Betz, S., Edlund, J., Gustafson, J., Eje Henter, G., et al. (2019). "Speech synthesis evaluation–state-of-the-art assessment and suggestion for a novel research program," in *Proc. 10th ISCA Workshop on Speech Synthesis (SSW 10)* (Vienna), 105–110. doi: 10.21437/SSW.2019-19

Weiss, B., Trouvain, J., Barkat-Defradas, M., and Ohala, J. (eds.). (2020). *Voice Attractiveness: Studies on Sexy, Likable, and Charismatic Speakers. Prosody, Phonology, and Phonetics*. Singapore: Springer. doi: 10.1007/978-981-15-6627-1

frontiers | Frontiers in Communication

TYPE Original Research
PUBLISHED 30 September 2022
DOI 10.3389/fcomm.2022.910952

Check for updates

OPEN ACCESS

EDITED BY
Sao Luis Castro,
University of Porto, Portugal

REVIEWED BY
Weimin Toh,
Nanyang Technological University, Singapore
Likan Zhan,
Beijing Language and Culture University, China

*CORRESPONDENCE
Ïo Valls-Ratés
io.valls@upf.edu

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Language Sciences, a section of the journal Frontiers in Communication

RECEIVED 01 April 2022
ACCEPTED 23 August 2022
PUBLISHED 30 September 2022

CITATION
Valls-Ratés Ï, Niebuhr O and Prieto P (2022) Unguided virtual-reality training can enhance the oral presentation skills of high-school students. *Front. Commun.* 7:910952. doi: 10.3389/fcomm.2022.910952

COPYRIGHT
© 2022 Valls-Ratés, Niebuhr and Prieto. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Unguided virtual-reality training can enhance the oral presentation skills of high-school students

Ïo Valls-Ratés [1*†], Oliver Niebuhr [2†] and Pilar Prieto [1,3†]

[1]Department of Translation and Language Sciences, Universitat Pompeu Fabra, Barcelona, Spain, [2]Centre for Industrial Electronics, University of Southern Denmark, Sønderborg, Denmark, [3]Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain

Public speaking is fundamental in our daily life, and it happens to be challenging for many people. Like all aspects of language, these skills should be encouraged early on in educational settings. However, the high number of students per class and the extensive curriculum both limit the possibilities of the training and, moreover, entail that students give short in-class presentations under great time pressure. Virtual Reality (VR) environments can help speakers and teachers meet these challenges and foster oral skills. This experimental study employs a between-subjects pre- and post-training design with four Catalan high-school student groups, a VR group ($N = 30$) and a Non-VR group ($N = 20$). Both groups gave a 2-min speech in front of a live audience before (pre-training) and after (post-training) 3 training sessions (one session per week) in which they practiced public speaking either in front of a VR audience or alone in a classroom (Non-VR). Students assessed their anxiety measures right before performing every speech and filled out a satisfaction questionnaire at the end. Pre- and post-training speeches were assessed by 15 raters, who analyzed the persuasiveness of the message and the charisma of the presenter. Speeches were also analyzed for prosodic features and gesture rate. First, results showed that self-assessed anxiety got significantly reduced at post-training for both conditions. Second, acoustic analyses of both groups' speeches show that the VR group has, unlike the Non-VR group, developed a more clear and resonant voice quality in the post-training speeches, in terms of higher cepstral-peak prominence (CPP) (although no significant differences in f0- related parameters as a function of training were obtained), as well as significantly less erosion effects than the Non-VR group. However, these differences across groups did not trigger a direct improvement on the participants' gesture rate, persuasiveness and charisma at post-training. Furthermore, students perceived the training in the VR environment to be more useful and beneficial for their future oral presentations than the Non-VR group. All in all, short unguided VR training sessions can help students feel less anxious, promote a more clear and resonant voice style, and can prevent them from experiencing an erosion effect while practicing speeches in front of a real audience.

KEYWORDS

public speaking skills, gesture, virtual reality, anxiety, charisma, prosody

Introduction

Boosting public speaking abilities in secondary school settings contributes not only to strengthening students' effectiveness with academic work (cf. the anecdote in Fox Cabane, 2013, pp. 139–141), but also their social skills, thus affording them more satisfactory interpersonal relationships (e.g., Morreale et al., 2000; Bailey, 2018) and preventing them from abandoning their studies prematurely (e.g., Boettcher et al., 2013; Niebuhr, 2021). In order to achieve these goals, it would be desirable that high schools acknowledge the importance of oral abilities for enhancing students' self-confidence and that they take action by involving students more often in oracy settings that encourage them to actively take part in their community (Bailey, 2018). However, time restrictions and the pandemic situation make it difficult for teachers to organize oral practices in front of the classroom. The present paper assesses the use of virtual reality technology (henceforth VR) as an alternative and complementary educational method for practicing oral presentations. Given the fact that VR can easily simulate traditional training scenarios in a virtual environment, the present investigation will determine the effects of a short 3-session VR training with high school students on reducing their public speaking anxiety and enhancing the quality of their oral presentations after training.

The importance of public speaking practice in educational settings

As any other skill, public speaking needs practice. One of the widely used instruction techniques in the educational system is the delivery of oral presentations by students, as they are frequently asked to present their projects or research papers in front of their peers. Yet one of the problems students face with this type of task is the fear of public speaking. PSA (or Public Speaking Anxiety, also called glossophobia) is related to different physiological changes like elevated heart and breathing rates, over-rapid reactions, trembling of muscles and shoulder and neck area stiffness (Tse, 2012). High levels of PSA can result in poor speech preparation (Daly et al., 1995) and impede decision-making of effective speech introduction strategies (Beatty and Clair, 1990; Beatty, 1998). Also, highly anxious individuals may be perceived by the audience as more nervous, they make less eye contact and pause more often than less anxious individuals (Daly et al., 1995; Choi et al., 2015); and most obviously, the quality of their speech performance is negatively affected (Beatty and Behnke, 1991; Menzel and Carrell, 1994; Brown and Morrissey, 2004). The negative thinking of those speakers exhibiting larger levels of PSA can reduce their speaking competence Daly et al., 1995; Rubin et al., 1997, and make them procrastinate in speech preparation (Behnke and Sawyer, 1999).

In practice, PSA and speech delivery problems can be effectively addressed by offering students more opportunities to rehearse their oral presentations. Goberman et al. (2011) showed that the earlier speakers started rehearsing their presentations on their own (i.e., unguided), the more fluent their speeches were after practicing, but with narrower pitch variation ranges compared with the students who started practicing later. A similar "prosodic erosion" effect (a successive lowering and narrowing of their speech melody across the repeated rehearsals of their presentation) is reported by Niebuhr and Michalsky (2018) (see also Niebuhr and Tegtmeier, 2019). Importantly, research shows that oral skills practice optimally needs to be performed orally in front of an audience. Smith and Frymier (2006) found that, compared with students rehearsing alone, rehearsing in front of an audience gave students higher scores on their final classroom-speech assessment, thus lending support to the claim that audience-based speech practice can help increase public speaking performance. Menzel and Carrell (1994) showed that practicing oral presentations before a classroom audience is the single greatest predictor of student speaking success and key for reducing PSA.

However, organizing such a setup can be difficult for teachers, given the high number of students per class and the extensive curriculum that needs to be covered in courses. The situation has been aggravated with the pandemic situation, where face-to-face interaction was limited to a great extent. Moreover, a high percentage of students dedicate most of their time to writing their speech rather than to rehearsing it orally, spending an average of <5 min on oral rehearsing (see Pearson et al., 2006). Given this situation, in the following section we assess the previous literature on the value of using VR as a complementary educational tool for providing an appealing setup for practicing audience-based oral presentations and thus boosting public speaking skills.

A complementary solution: Empirical evidence on the effects of VR for boosting public speaking skills

As a way to enhance the oral practice of presentations and, also, to reduce anxiety when delivering speeches in front of an audience, VR simulations can be of great help. Virtual simulations can be broadly defined as 3D interactive environments that are computer-generated and are viewed by a single user through a headset that excludes all other visual input. While many of these VR platforms have been traditionally used for entertainment purposes, a large number of schools, hospitals, and research institutions (Peeters, 2019) are currently using this technology to provide active learning environments

(Legault et al., 2019). Since VR experiences evoke realistic responses in people, they can be fundamentally conceived as "reality simulators." Participants in VR settings are placed in an artificial scenario that depicts potentially real events, with the likelihood that they will act and respond realistically. VR gives rise to the subjective illusion that is referred to in the literature as *presence*—the illusion of "being there" in the environment depicted by the VR displays—in spite of the fact that the user is simultaneously fully aware that the environment is artificial (Armel and Ramachandran, 2003). VR is different from other forms of human–computer interface "since the human participates in the virtual world rather than uses it" (Slater and Sanchez-Vives, 2003, p. 3). Mikropoulos and Natsis's (2011) empirical study dealing with the application of virtual reality in learning environments suggests that "presence is considered to be a key feature" with a majority of the practitioners, whose work they examined, reporting that "their sample had the feeling of 'being there' and that this might contribute to positive results" (p. 774). Accordingly, "being there" leads to the participants' increase in "intrinsic motivation and engagement" (Dalgarno and Lee, 2010). Ruscella (2019) and LeFebvre et al. (2020) suggest that an immersive setting reduces fear and creates a no-risk situation that is ideal for learners to practice their speeches. As LeFebvre et al. (2020, p. 10) points out, "VR creates a more effective treatment environment for enacting changes to reduce PSA." Even though information about public speaking might not be provided to the user, spending time practicing in front of the virtual audience may improve social skills that can be transferred to the real world (Xu et al., 2011; Lane et al., 2013; Rogers et al., 2017; Howard and Gutworth, 2020).

Effects of VR to treat public speaking anxiety

In the context of public speaking training, some studies have tested the use of VR technology to reduce anxiety in university students. In a systematic review, Daniels et al. (2020) identified 14 studies conducted from 2009 to 2019 that used VR as a tool to diminish public speaking anxiety (PSA). From these 14 studies, 7 belonged to clinical settings (Wallach et al., 2009, 2011; Lister et al., 2010; Lister, 2016; Lindner et al., 2018; Yuen et al., 2019; Zacarin et al., 2019). Three of the 7 clinical studies (Wallach et al., 2009, 2011; Lister et al., 2010) compared PSA levels before and after VR immersion and found a significant PSA reduction. Wallach et al. (2009) compared, with 88 participants, Cognitive Behavioral Therapy (CBT) to VR immersion in a total of 7 sessions, and they found that both treatments were effective in reducing speakers' anxiety (see also (Safir et al., 2012)). In a later study, Wallach et al. (2011) applied the same design, with 20 female participants, this time comparing Cognitive Therapy (CT) to VR. They yielded the same results regarding both treatments. Lister et al. (2010), in a study with 20 participants, found that VR 3D videos were capable of eliciting a fear response in participants and was effective in reducing negative self-beliefs about public speaking abilities. In the study by Lindner et al. (2018), with 50 participants, they compared therapist-led exposure followed by 4 VR internet intervention sessions to a self-led waiting list (WL) condition. They concluded that those internet interventions were as effective as the traditional therapist-led interventions in reducing speakers' PSA. Moreover, VR intervention sessions showed that this cost-effective technology can lead to solid and promising automated self-help applications. In another study by Lindner et al. (2020) with 25 participants, they showed that only one session of VR exposure therapy constituted an effective treatment of PSA. Lister (2016), in a study with 98 participants that compared a VR condition to a control condition, concluded that six sessions were capable of increasing confidence of speakers and obtained positive self-statements. Two clinical studies included in the systematic review did not include control conditions, namely Yuen et al. (2019) and Zacarin et al. (2019). Yuen et al. (2019) in two pilot studies with 11 and 15 participants each, showed that 6 weekly sessions were enough to significantly reduce PSA in a 3-month follow-up test. In the study by Zacarin et al. (2019), with 6 female participants, they designed 6 individual sessions and 1- and 3-month follow-up sessions, all including feedback by the therapist. Results showed that feedback allowed them to improve their speech and that this contributed to reducing their anxiety. Also, an increase in speaking quality was found in terms of a reduction of silent pauses and of word repetitions.

The other 7 studies included in the systematic review were performed in university educational settings. Two of them compared PSA from pre to post treatment and found a significant reduction (Heuett and Heuett, 2011; Nazligul et al., 2017), whereas the other five had different research designs. Heuett and Heuett (2011) carried out a study with 80 university students. The pre-training sample gave an impromptu speech and filled out questionnaires related to PSA and Willingness to Communicate (WTC)—and was then randomly assigned to one of three groups. One group practiced public speaking to a VR-generated virtual audience, another group was trained to visualize an audience as they spoke, and the third group, i.e., the control group, received no training at all. Both treatments lasted between 10 and 20 min, after which all three groups carried out a post-test which was identical to the pretest, and all participants completed the same questionnaires again. A comparison of pre-training and post-test data from the participants in the VR group showed a significant reduction in trait and state communicative apprehension (CA), and an increase in their self-perceived communication competence (SPCC) and WTC scores. The visualization treatment also yielded significant improvements in trait and state CA and SPCC, but not in WTC. The control group reported no significant change for any of the variables studied. The other study, by Nazligul et al. (2017), was conducted with 6 software engineers university students (21 years old). Every participant attended a 1-h individual therapy session

where they were told about anxiety and its possible causes and components, and they rated their self-perceived anxiety level while imagining giving a speech. After that, they performed a brief speech on a controversial topic and rated their self-assessed anxiety with the SUDS at 4 different points during exposure. Participants reported that, while being exposed to VR, they felt the highest level of anxiety, but also lower levels of anxiety after the intervention ended. There was no control group.

Two other educational studies that had no control group were Stupar-Rutenfrans et al. (2017) and Takac et al. (2019). The former was conducted with 19 university students and demonstrated in a within-subject task design that rapidly successive VR scenarios could elicit self-reported distress, and significant physiological arousal was also observed in heart rate data. Distress was easier to trigger than habituation, with three successive speeches (within a 60-min session) required to sustain distress reduction. Stupar-Rutenfrans et al. (2017) carried out a further study in which 35 university students performed three different speeches, one per week, using VR technology at home. In the first session the VR screen showed no audience, in the second the VR screen showed a small audience and in the third, a large audience. Participants had to fill out three questionnaires to assess their levels of anxiety and emotion regulation during treatment: namely the Emotion Regulation Questionnaire (Gross and John, 2003), the Public Report of Communication Apprehension (McCroskey, 1982), and the STAI Inventory. The study concluded that initially more anxious participants significantly improved in self-assessed anxiety scores after having performed in all three VR conditions. Their anxiety increased between the first and second session but diminished before and after the third session. The authors recommended that future research in that line should include a control group and also pre- and post-training tasks that would include speaking to a live audience in order to compare the reduction of anxiety in virtual and non-virtual public speaking contexts.

Aymerich-Franch and Bailenson (2014), North et al. (2015), and Wilsdon and Fullwood (2017) conducted educational studies that included both a VR and a control condition. The former study had a total of 14 participants and compared VR (7 participants) to a no-treatment group (7 participants) in a total of 5 sessions. They found a significant reduction in fear measures in the treatment group, but no relative comparison between groups was made. Aymerich-Franch and Bailenson (2014), with a sample of 41 participants, conducted a study with a VR group that performed visualization with a doppelganger (virtual humans that highly resemble the real self but behave independently) and a control condition that performed visualization with imagination. For VR participants, the first part of the session consisted of seeing their doppelganger performing a successful speech through VR while listening to a relaxing voice. The control group had to imagine giving a successful speech while listening to the relaxing voice. After that, participants of both groups performed a speech on a topic of their choice before an audience of two people. They concluded that there were no differences in self-perceived anxiety across groups. However, they found an interaction between condition and gender for state anxiety and self-perceived communicative competence. The doppelganger technique worked better for males, and as the authors point out this was probably because men were already more familiarized to be in virtual environments and felt more comfortable during the VR experience, whereas the visualization technique proved more effective for females. To our knowledge, only one study has reported null effects of VR training on anxiety. Wilsdon and Fullwood (2017) conducted a one-session study with 40 university students consisting of 3 VR conditions (high, medium, and low immersion environments) and a control condition. The VR conditions performed a 5-min speech about their first week at university before a VR audience, while the control condition performed the same speech to the researcher. Participants filled in anxiety self-assessment questionnaires before and after the speech task. Results showed no improvement in PSA reduction, and increased VR immersion did not significantly reduce their anxiety either.

Besides the studies included in the systematic review, there are other studies that also show positive results in anxiety reduction: Harris et al. (2002) in a study that involved 14 university students with a VR group and a WL group, found that four 15-min sessions of VR were effective for reducing PSA. The pre-training consisted of different short public speaking tasks and different self-report instruments. The VR group then underwent four training sessions with different tasks while the WL group was given the same VR training once the experimental data had been gathered. Post-testing consisted of the same respective tasks. Although there were significant reductions in anxiety at post-test on some measures in the VR group (self-assessed questionnaires and heart rate), only one comparison between the VR and the WL group proved to be significant—i.e., the one that compared levels of speaker self-confidence. VR participants showed greater improvement overall on both self-assessment and physiological measures. Rodero and Larrea (2022) conducted a study with 100 university students, and they were divided into a VR experimental group and a control group. They performed a pre-training and a post-training task which consisted of giving a 3-min speech in front of a live audience. Trainings consisted of 5 trial sessions with a VR environment for the experimental group, whereas for the control group the 5 training sessions were led by an instructor. During the training sessions in both conditions, the authors included distractors (someone coughing in the audience or someone in the audience asking a question). The study measured self-assessed anxiety and electrodermal activity. Results show that VR participants significantly reduced their anxiety levels (in both measures) and that distractors (someone coughing placed at second 40 and someone's question at second 60, in pre- and

post-test speeches) proved effective at reducing their anxiety at post-test. Therefore, they conclude that training with distractors is effective and reproduces a more real public speaking situation. Participants said that training with VR helped them concentrate, made them more confident and made them have less tension.

To our knowledge, only one study (Kahlon et al., 2019) has previously examined VR effects on PSA reduction in a secondary school setting. They studied the PSA of 27 adolescents (aged 13–16) after only a single 90-min VR session, in which they performed different speaking or public speaking exercises. Subsequently, they received brief psychoeducation, active maintenance and filled in different anxiety self-assessments. A therapist accompanied them throughout the session. The authors concluded that one session was enough to reduce PSA of adolescents after 1- and 3-month follow ups, although the causes for this PSA reduction are not clear as there were neither control nor comparison groups.

Effects of VR on students' motivation

All in all, there is evidence that VR serves as a tool to trigger anxiety during training sessions and eventually reducing anxiety after training. However, in the context of educational practice, are VR public speaking trainings capable of stimulating a higher commitment to learning, in particular with respect to high-school students as the target group?

Several studies have shown that students are highly motivated using VR technology for practicing public speaking. The study by Frisby et al. (2020) concludes that employing VR for speech rehearsals not only helps diminish PSA. Rather, students consider it an innovative way of oral rehearsing that makes them more willing to accomplish a good performance. Vallade et al. (2020) and Kryston et al. (2021) also report on the excitement of students to participate in VR experiments as a different and motivating way to entice them to rehearse their speeches. Specifically, Kryston et al. showed how participants in the VR settings reported that it was more demanding than other modes of practice, which is consistent with the ability of digital audiences to elicit mental stress in speakers. In their qualitative study, Gruber and Kaplan-Rakowski (2020) examined the efficacy of VR based on the perception of 12 university students performing 8 different speeches. They analyzed the participants' sense of presence, the plausibility of the illusion and the perceived usefulness of VR for practicing public speaking. Although the sample was small, participants acknowledged the potential of VR for practicing oral speeches, compared to traditional practices, they saw cognitive benefits of the VR experience and they would find it useful as a tool to practice oral presentations to be presented in front of university audiences. They also emphasized how practicing with VR made them more capable of speaking in front of live audiences. Findings by Daniels (2021) showed that the usability ratings of virtual reality as a training tool for public speaking training can vary depending on the technological background of users. They concluded that "the use of virtual reality as a training tool for public speaking training is highly recommended. This is supported by the unanimously positive responses of participants in the System Usability Scale (SUS) that measures their interest in using the VR tool for oral presentations" (Daniels, 2021, p. 6).

Effects of VR as conducive of a more listener-oriented prosodic style

Given that VR provides a credible set of scenarios that allow for an immersive learning situation, when used for public speaking tasks, VR environments have been reported to be conducive to a more listener-oriented speaking style from the point of view of the prosodic characteristics. To our knowledge, five studies have assessed the impact of using VR on the speech characteristics of the speakers while using this technique during a public speaking task as compared to other conditions. Three of them (Niebuhr and Michalsky, 2018; Remacle et al., 2021; Valls-Ratés et al., 2021) put the focus on prosody (which refers to all aspects of a speaker's voice and tone-of-voice). Niebuhr and Michalsky (2018) showed in a study with 24 participants comparing VR and Non-VR groups, that those students rehearsing public speeches within a VR environment performed their speech in a more listener-oriented, conversation-like speaking style than participants in the control group, who practiced their speech alone in a classroom. They concluded that the speeches of participants who were trained in the VR condition were more charismatic and more audience-oriented (characterized by a higher F0 level, a larger F0 range, and a slower speaking rate), showing reduced signs of "prosodic erosion" due to repeated rehearsing, compared to those participants who had practiced their speeches alone in a classroom (see also Niebuhr and Tegtmeier, 2019). Moreover, compared to the control Non-VR group, the speakers were unexpectedly motivated to speak longer, and the speech of the VR group was characterized by higher fundamental-frequency (i.e., f0) levels, a wider f0 range, a slower speaking rate, fewer pauses and a higher intensity level. A recent study by Remacle et al. (2021) conducted with 30 female elementary school teachers also proved to be effective in prompting vocal characteristics that are very similar to the ones used in the classroom. Teachers gave the same lesson in their classrooms and later in front of a VR audience. Results showed that, in line with Niebuhr and Michalsky (2018), performing both in front of real and virtual audiences (compared to free speech performed before the experimenter in a control condition) significantly increased the participants' f0 values, their f0 variations and their voice intensity levels. Another recent study by Valls-Ratés et al. (2021) utilizing the same corpus used in the present study, with 31 participants, found that VR trainings induced a more audience-oriented prosody, making participants increase their f0 values, they spoke for longer time, there was an increase

in the number of pauses, and they also increased their gesture rate throughout the VR sessions. A study by Notaro et al. (2021) analyzed the effects of VR on fluency and gesture rate after 13 participants (20–25 years old) performed the same speech at two different times: the first time in front of a real audience and the second time in front of a VR audience, while also having the same real audience in front of them. They analyzed vocal parameters during VR and audience-based training and concluded that participants had a higher voice modulation, more voice power and paused more often when using VR. They also lowered their speech rate as well as their number of gestures per minute, pointing to the possibility that there existed a higher control over gestures while speaking with the VR glasses on. Finally, focusing on an L2 setting, Thrasher (2022) conducted a study with 25 participants (22 years old, L2 learners of French) that lasted 9 weeks. In order to assess the L2 speech in VR and Non-VR contexts, participants were asked to perform four public speaking tasks, two VR tasks and two in-class tasks. When French raters assessed the audio files, they found that the speech of participants using VR was more comprehensible than the speech of participants performing in-class.

Given that the studies reported in this section have shown that using VR for public speaking tasks triggers a more listener-oriented speech style, it is plausible to expect that a VR-training paradigm will trigger a more audience-oriented speech style in post-training speaking tasks. Yet to our knowledge very few studies have assessed the effects of VR on public speaking performance (see the next section).

Effects of VR on public speaking performance after training

To our knowledge, only two studies have been conducted to assess the effectiveness of VR public speaking training on public speaking performance after training. In a recent study, Sakib et al. (2019) performed a 3-month VR public speaking training study with a pre- and post-test design with 26 participants. Pre- and post-training speeches were performed in front of a real audience, whereas treatment consisted of 8 sessions in front of VR audiences. They collected a variety of measures of self-assessed and physiological anxiety, as well as ratings on speech performance assessed by external raters using an assessment form to rank speaker's performance from 1 (highest score) to 5 (lowest score). Results showed that participants improved their public speaking performance from pre- to post-training and also significantly reduced their self-assessed anxiety indicators, as well as two physiological anxiety measures (skin conductance response and skin temperature), resulting in a match between self-assessed and physiological markers. Even though the study concluded that VR environments were effective in reducing speakers' anxiety and enhancing public speaking performance, there was no control group to compare these results to and public speaking performance was assessed in general terms. The second between-subject study by Van Ginkel et al. (2020) compared general public-speaking performances before and after VR public speaking training by involving both a VR and a Non-VR control group. The authors conducted a VR training study with 22 pre-university students across a 2-week period that consisted of three sessions: in the first and third sessions participants were introduced to the different features that an effective speech should include and after the instruction they had to give a 5-min speech in front of their peers. The second session was dedicated to performing a 5-min speech within a virtual environment, after which in a follow-up third session the VR condition received computer-mediated automatic immediate feedback and the control condition received delayed feedback given by an expert. The authors concluded that the VR session together with the given feedback was effective in improving eye contact and pace when delivering a speech in front of a real audience. However, they also pointed out that it is difficult to claim that the results are a direct consequence of the VR practice itself, as the instructions given to them, the feedback, and the independent practice could have had an influence as well.

Interestingly, in an L2 language learning context, Gao (2022) conducted an 8-week public speaking training study in which 90 Chinese university students participated in either a VR condition or a control condition based on traditional multimedia technology to test their proficiency in spoken English. After 8 weeks of autonomous learning, students were tested at post-training with English reading materials and oral presentation of specific topics. While participants in both conditions were successful in improving the oral English pronunciation skills (in this study they add the role of speech emotion to the usual pronunciation assessment systems that consider only the tone, intonation and rhythm of speech), the VR condition outperformed the control condition.

All in all, the investigations assessing the value of public speaking VR training initially point out to a gain in public speaking performance in terms of general performance, eye gaze and speech rate (Sakib et al., 2019; Van Ginkel et al., 2020). Importantly several studies have indicated that VR triggers a more listener-oriented speech style (Niebuhr and Michalsky, 2018; Niebuhr and Tegtmeier, 2019; Notaro et al., 2021; Remacle et al., 2021; Valls-Ratés et al., 2021). Yet to our knowledge no previous investigation has assessed the value of VR training by assessing public speaking performance at post-test by incorporating a full-fledged prosodic analysis of the post-test speeches. We expect that the observed effect of VR in triggering an audience-oriented speech style will also carry over into the speakers' post-training speeches.

The present study: Main goal and hypotheses

Against the outlined research background, still very little is known about the potential boosting effects of practicing oral presentations with VR on developing students' public speaking skills and whether the training has an impact on the prosodic and gestural characteristics of the post-test speeches. Therefore, the main goal of this study is to investigate, through a between-subjects training experiment, whether training in public speaking with VR environments makes a difference in the overall quality of the oral presentations that students perform in front of an audience after training. To our knowledge, this is the first VR public speaking training experiment conducted with high school students that investigates not only the effects of training with VR on self-perceived anxiety both in the pre- and a post-training public speaking tasks but also on overall public speaking performance (through the use of persuasiveness and charisma ratings), as well as on oral presentation quality through a systematic analysis of the prosodic and gestural features of those oral presentations. Importantly, the assessment of the two speeches given in front of a live audience, e.g., before and after training, will be comprehensive. First, we will assess how the speaker feels in terms of self-perceived anxiety. Second, we will also include assessments about the persuasiveness of the speakers' charisma by external raters that are blind to the conditions. In addition, we will assess the prosodic characteristics of these speeches (understood holistically as involving a set of parameters including f0, tempo and voice quality characteristics), as well as the gesture rate, and the level of participants' own satisfaction after the training.

The following hypotheses will be tested: (a) Compared to the Non-VR public speech training, VR-based speech training will help diminish public speaking anxiety in the post-training public speaking task in front of a real audience. (b) VR public speaking training will lead to higher persuasion and charisma ratings. (c) VR public speaking training results in prosodic differences compared to the baseline condition of speakers, making the resulting speech more audience-oriented. (d) The audience-oriented prosody will be associated with a higher number of gestures in the VR condition. (e) Participants of the VR condition find more enjoyment and report a higher motivation for their future oral presentations.

In sum, the purpose of this educational intervention was to examine the impact of VR public speaking training on the quality of public speeches performed after training in front of a live audience, by comparing it to a Non-VR condition in which speeches were rehearsed individually. An important component of this assessment includes a complete analysis of the prosodic features of these speeches. In this way, we assess the value of a complementary use of a VR tool that can help educators promote the rehearsal of oral presentations and ultimately improve students' oral skills.

Methods

We designed a between-subjects training experiment with a pre- and post-test experimental framework. The public speaking training involved three training sessions, one per week (three for the VR condition and three for the Non-VR condition). Both before and after the training, a public speaking task was performed individually in front of a real audience, see Figure 1. The total duration of the experiment, from the pre-training to the post-training public speaking task was 5 weeks.

Participants

A total of 65 secondary school students aged 17–18 were recruited from four high schools (Institut Fort Pius, Institut Quatre Cantons, Institut Vila de Gràcia and Institut Icària) in the Barcelona area. These high schools are located in two central city quarters of Barcelona. The study was supported by the four school boards, which treated the proposed training as an extra-curricular activity which was carried out in the school premises. These four high-schools were chosen because they are placed in two central districts of Barcelona (Gràcia and Sant Martí), with very similar Catalan-Spanish language dominance (the percentage of Catalan speaking students being 81.9 and 78.8%, respectively), and with similar middle-income social composition.[1]

Of the original 65 participants, 14 participants' data had to be disregarded for one of the following two reasons, namely (a) because of participants being absent at one of the training sessions or at the post-training phase, or (b) because their speeches at either pre- or post-test did not reach the minimum duration that we established (i.e., 1 min) or because they did not offer a minimum of two arguments to support their persuasive speech. The 50 remaining participants (mean age = 16.95, SD = 0.17; 70% female and 30% male) completed all five speeches with the required characteristics. Participants were randomly assigned to either the VR group ($N = 30$) or the Non-VR group ($N = 20$).

All participants were typically developing adolescents and had no history of speech, language, or hearing difficulties. Participation was voluntary, and all participants completed an informed consent form during the initial training session. Participants performed their speeches in Catalan. All students were bilingual Catalan-Spanish speakers, with 89.7% of them naming Catalan as their dominant language. The main language of instruction in the target schools is Catalan.

1 Anuaris Estadístics de la Ciutat de Barcelona. 1996–2020 (*Barcelona's Statistical Annual Directory*): https://ajuntament.barcelona.cat/estadistica/catala/Anuaris/Anuaris/anuari19/cap06/C0616010.htm.

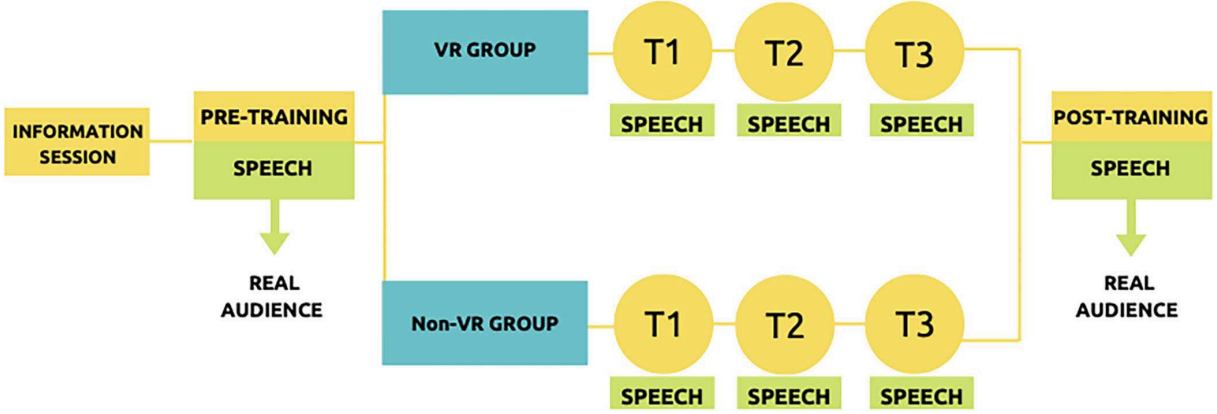


FIGURE 1
Experimental design.

Materials for the public speaking tasks

A total of 5 short public speaking tasks had to be performed individually by each participant, two in front of a real audience (i.e., the pre-training and the post-training public speaking tasks), and three for training purposes. For all the public speaking tasks, participants were given a specific topic and a sheet of instructions (see Appendix) containing a list of arguments they could use in order to prepare a persuasive speech. In all cases, they were asked to prepare a 2-min speech.

An initial choice of 10 topics was first made based on a long list of suggested topics taken from a website maintained by instructors of public speaking and other communication courses (i.e., www.myspeechclass.com). This initial list of 10 topics was assessed through an online questionnaire which was distributed to mailing lists of 17-year-old boys and girls. A total of 58 anonymous students participated in the poll. They were asked to vote on their favorite topics from 1 (least liked) to 7 (most liked). The topic selected for both pre-training and post-training public speaking tasks was the same, namely: "Do you think that adolescents should spend more time in nature?". In order to minimize the argumentation and expression differences across participants, five possible arguments were provided to participants. They were also given 2 min to prepare their speech. Though they could take notes for that purpose if they wished, they were not allowed to use the notes when they delivered their speech to prevent them from reading the whole speech.

The three topics for each of the three VR and Non-VR training sessions were the following: "What would the house of my dreams be like?", "Is graffiti a form of art?", and "Can happiness be bought?". The instructions given to participants for the preparation of their speeches during the training sessions was the same as the instructions given to them for the pre- and post-test public speaking tasks.

Experimental design

The structure of this between-subjects training study was a pre-training phase followed by a training period and a post-training phase (see Figure 1). One week prior to the pre-training phase, an information session was organized by the experimenter in each of the high-schools and served the purpose of preparing the students for the pre-training session and explaining the experiment's procedure and overall schedule that participants would have to bear in mind when delivering a speech. Pre and post-training sessions were also conducted by the experimenter and a research assistant. Both the research assistant and the 3-people live audience were blinded to the procedure of the study. During the information session participants were instructed on how to use VR and they could familiarize themselves with the VR goggles.

They were told that an audience of three people would attend their speech. They also knew that the pre-training speech would have to be persuasive, and that it was to be performed to convince three representatives of the Catalan Government to take action. Yet the topic itself would only be revealed to them immediately before the speech. After this, each group of students was randomly divided into the VR and the Non-VR group. The VR group performed the three training sessions delivering their speeches in front of a virtual audience, whereas the Non-VR group gave the same set of speeches while being alone in a classroom. The reason to choose three short VR sessions was based on the belief that adaptation to the virtual context would need some repetitions. Empirical reports of fast and reliable learning of visual context-target associations have proved effective after just three repetitions (Zellin et al., 2014). Finally, all participants carried out a post-training, which consisted of the same persuasive public speaking task as the pre-training.

10.3389/fcomm.2022.910952

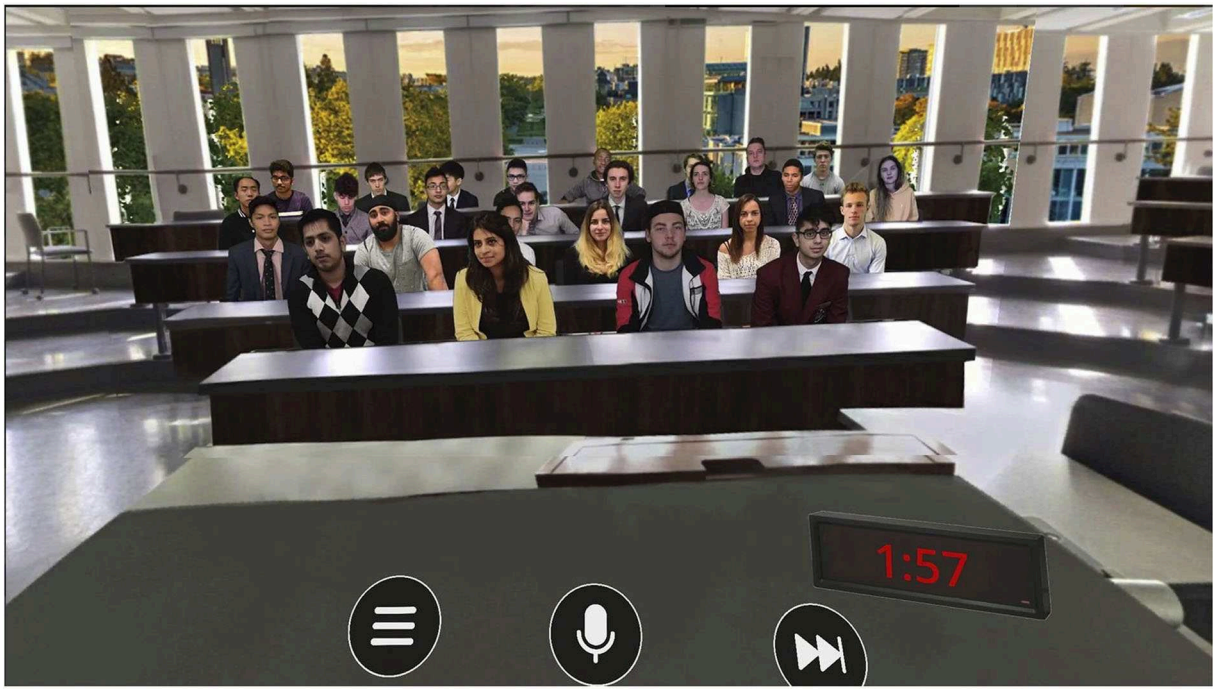


FIGURE 2
Screenshot of the VR scenario with a virtual audience generated by BeyondVR.

In order to pilot the materials, topics and procedure of the experiment, four 17-year-old students participated in a 3-h pilot session in which they were asked to prepare 3 speeches in 2 min to give in front of a small audience following our target set of instructions. The instructions informed participants of the amount of time they would have to prepare and to deliver the speech. For every speech they were given a written script of ideas related to the topic that they could use to include in their presentations. The pilot session contributed to refine and validate the final scripts and the procedure. For example, we realized that if speakers were allowed to use their written outline while speaking, they were reading from it most of the time. Therefore, we did not allow participants to have the outline with them to prevent them from reading and to enhance their connection with the audience.

Procedure

The experiment was performed individually in separate classrooms at the four high schools. The first author of the study was the experimenter and in charge of the data collection. All 5 public speaking tasks per student (3 during the training phase and 2 at pre- and post-training) were video recorded.

All participants started with the same pre-training task, which consisted of giving a brief speech in front of a live audience. Before giving their speech, participants received a sheet of instructions in which they were asked to prepare and then deliver a 2-min persuasive speech in front of three representatives of the Catalan Department of Education to convince them to increase funding for secondary school field trips to the countryside. Participants were allotted 2 min to prepare their speech and did so alone in an empty classroom. After the 2 min of preparation had elapsed, they went to the adjacent classroom. The procedure was repeated for the post-training public speaking task.

For the training sessions, the procedure was largely similar between the two conditions. The Non-VR participants entered the classroom and were given the instructions. When they felt ready, they started performing the speech, with a visible timer that counted down the 2-min speaking time for them. For the VR participants, the only difference to the Non-VR participants was that right before practicing the speech, the experimenter fitted them with a Clip Sonic® VR headset to which a smartphone was attached. Using the free BeyondVR virtual reality interface application installed on the smartphone, the VR headset created the 3D illusion that the participant was standing in front of an audience. The virtual audience in this application moves while sitting and they show a sympathetic stance while the participant is speaking. They all look at the speaker and show interest in what the speaker is talking about, see Figure 2. Note that a timer is also visible in the view provided

by the VR headset to allow speakers to monitor their use of time and not exceed the 2-min limit. Although we did not control for previous use of VR among participants, none reported any kind of discomfort wearing the VR goggles.

Anxiety measures

In order to control for anxiety and to facilitate comparisons with studies that have assessed anxiety in public speaking tasks through self-perception measures we used, as well as previous studies (e.g., Macinnis et al., 2010; Heuett and Heuett, 2011; Verano-Tacoronte and Bolívar-Cruz, 2015), the Subjective Units of Distress Scale, henceforth SUDS (Wolpe, 1990), a validated and widely used self-assessed anxiety questionnaire which uses a 100-point scale anchored on 0 (no fear), 25 (mild fear), 50 (moderate fear), 75 (severe fear), and 100 (very severe fear). Subjective distress refers to uncomfortable or painful emotions felt, and thus SUDS is used to systematically gauge the level of distress. The SUDS scale was developed by Wolpe (1969) and has been frequently used in Cognitive Behavioral Therapy (CBT) to evaluate treatment progress. Participants were given the SUDS assessment sheet just prior to entering the room where they would give their pre- and post-training speeches.

Satisfaction questionnaire

One month after the experiment ended, a brief online satisfaction questionnaire was sent to all participants asking the following three questions: "Did you feel comfortable participating in the experiment?", "Did you have fun?" and "Did you find the experiment useful for your current oral presentations?". They were asked to assess their satisfaction level using a Likert scale that ranged from 1 to 10. Nine (out of the 20) Non-VR participants and 19 (out of the 30) VR participants answered the online survey.

Data analysis

A total of 100 pre-training and post-training speeches were obtained from the 50 participants (50 participants × 2 pre- and post-training speeches). The target persuasive speeches were assessed for the following features, namely (a) persuasiveness and charisma (Persuasiveness and charisma); (b) voice parameters (Voice parameters); (c) manual gesture rate (Manual gesture rate) and a satisfaction questionnaire (Satisfaction questionnaire). Apart from these measures on the actual speeches, a self-perceived anxiety SUDS measure and the results of the satisfaction questionnaire were also included in the data analysis.

TABLE 1 Survey questions regarding persuasiveness and charisma.

From 1 to 7 rate how persuasive the message is.	From 1 to 7 rate how charismatic the speaker is.

Persuasiveness and charisma

In order to assess the persuasiveness of pre- and post-training speeches, as well as the charismatic value of the speaker, a group of 15 raters (9 women and 6 men) with an age range from 23 to 63 years carried out a rating task on the speakers' persuasiveness and charisma, based on the video recordings of each presentation. The raters were chosen such that all had a university degree and that, overall, the rater sample was balanced with respect to gender. A 1-h training session was held with all raters and the first author of the study, in which they were given instructions as well as some time to practice and familiarize themselves with their task. They were first offered definitions of persuasiveness [understood by Rocklage et al. (2018, p. 751) as: "deliberate attempt to change the thoughts, feelings, or behavior of others"] and charisma [taking the definition by Niebuhr et al. (2020) "communication style signaling leadership qualities such as commitment, confidence, and competence that affect followers' beliefs and behaviors in terms of motivation, inspiration, and trust"]. Raters were asked to watch each video recording and then provide responses to the three questions in Table 1. They were asked to assess persuasiveness and charisma of the speaker in an intuitive way, without carefully analyzing vocabulary nor rhetorical strategies. They were asked to rate the speeches as if they were watching TV, assessing from 1 to 7 how persuasive the message was and how charismatic they perceived the speaker was.

An online survey sheet with the questions in Table 1 was prepared using Alchemer[2] (formerly SurveyGizmo, 2006). The 100 speeches were distributed across four surveys to offer the raters enough time to have a break after each block of about 15 stimuli. The speeches were presented in pairs. Each pair consisted of either pre- or post-training speeches of the same speaker so that raters could listen to them one after the other and assess which of the two was better. The rating task for all the speeches took about 5.5 h. The raters received a monetary compensation of 10 EUR per hour. The inter-reliability score (ICC) was excellent 0.913 (i.e., results are considered reliable as the score exceeded 0.7) (Koo and Li, 2016).

Voice parameters

For each participant, the total durations of the recorded speeches were similar in the pre- and post-training conditions (M = 1:23 min; span = 1:00–2:00 min). The acoustic analysis included a total of 16 different vocal parameters (5 f0 parameters,

2 https://www.alchemer.com/

4 duration parameters, and 7 voice parameters; see below). The acoustic-phonetic analysis was automatically performed using the ProsodyPro script of Xu (2013) and the supplementary analysis script of De Jong and Wempe (2009), both with the (gender-specific) default settings of PRAAT (Boersma and Weenink, 2007).

In the f0 domain, we measured f0 minimum and maximum, the f0 variability (in terms of the standard deviation), the mean f0 and the f0 range. For all five f0 parameters, one value was determined per prosodic phrase. Measured values were checked manually for plausibility. Outliers or missing values were corrected by manual measurements. Moreover, all f0 values were recalculated from Hz to semitones (st) relative to a base value of 100 Hz. The prosodic domain of calculation for those f0 values was the interpausal unit (IPU), which was automatically detected. The criterion was the detection of an IPU boundary was the presence of a silent gap interval >= 200 ms, with silent gap being defined as a drop in intensity > 25 dB.

The tempo domain consisted of the following seven measured parameters: total number of syllables, total number of silent pauses (>300 ms, which is above the perceived disfluency threshold in continuous speech) (Lövgren and Doorn, 2005), total time of the presentation (including silences), total speaking time (excluding silences), the speech rate (syllables per second including pauses), the net syllable rate (or articulation rate, i.e., syll/s excluding pauses) as well as ASD, i.e., the average syllable duration. ASD is a parameter that closely correlates with the fluency of speech (Rasipuram et al., 2016; Spring et al., 2019). As De Jong and Wempe (2009) summarize in their literature review: "An advantage of using inverse articulation rate [ASD] is that [...] it is a measure of disfluency, in the sense that higher values (longer mean syllable times) mean less fluent speech" (p. 900). All temporal measurements were conducted based on the analyzed presentation as a whole.

The domain of voice quality measurements included the nine parameters that are very frequently used in phonetic research (e.g., for analyzing emotional or expressive speech, see Banse and Scherer, 1996; Liu and Xu, 2014): harmonic-amplitude difference (f0 corrected, i.e., h1*-h2*), cepstral peak prominence (CPP), harmonicity (HNR), h1-A3, spectral center of gravity (CoG), formant dispersion (F1–F3), median pitch, jitter,[3] and shimmer. Like for the f0 parameters, voice-quality measurements were conducted based on the prosodic phrase, i.e., one value per prosodic phrase was calculated. Also, all values were manually checked and corrected, if required. This meant that a trained phonetician conducted a visual inspection of the measurement tables and marked potential outliers, i.e., in particular, unplausible values such as "0 Hz" or "600 Hz" for mean f0 and f0 maximum or a F1–F3 formant dispersion of "−1 Hz", etc. these were corrected my manual re-measurements (or deleted from the dataset).

Manual gesture rate

First, all communicative gestures were annotated by taking into account the gestural stroke (the most effortful part of the gesture that usually constitutes its semantic unit; (McNeill, 1992; Kendon, 2004)). Non-communicative body movements (self-adaptors, e.g., scratching, touching one's hair; (Ekman and Friesen, 1969)) were excluded. Gesture rate was calculated per every speech as the number of gestures produced per speech relative to the phonation time in minutes (gestures/phonation time).

Satisfaction questionnaire

The means for each of the three questions of the satisfaction questionnaire and the reliability of the questionnaire (using Cronbach's Alpha) were calculated.

Statistical analyses

The statistical analyses were performed using IBM SPSS Statistics 19. A set of GLMMs were run for five independent variables, namely SUDS (anxiety), Persuasion and Charisma, Voice and Gesture rate. The models include Condition (two levels: VR and Non-VR) and Time (two levels: Time 1-pre-training; Time 2-post-training) and their interactions as fixed factors. Subject was set as a random factor. Pairwise comparisons and *post*-hoc tests were carried out for the significant main effects and interactions.

For the satisfaction results, an independent a *t*-test was performed for each of the three questions in the satisfaction questionnaire. To make sure that there was rater interreliability, we performed a Reliability Analysis using the Intraclass Correlation Coefficient (ICC).

Results

Self-assessed anxiety SUDS

The GLMM analysis for SUDS showed a main effect of Condition [$F_{(1, 96)} = 8.785$, $p = 0.004$], which indicated that in general (both at pre- and post-training) Non-VR values were higher than VR values ($\beta = 13.792$, SE = 4.653, $p = 0.004$), and a main effect of Time [$F_{(1, 96)} = 10.807$, $p = 0.001$], showing that SUDS values where lower at post-training regardless of the condition ($\beta = 8.292$, SE = 2.522, $p = 0.001$). No significant interaction between Condition and Time was obtained, showing that the two conditions were not significantly

3 "The term jitter describes the small period-to-period variation in f0 and hence deviation of a speaker's voice from strict periodicity" (Niebuhr et al., 2020, p. 13).

TABLE 2 Summary of the GLMM analyses for the 5 f0 variables, in terms of main effects and interactions.

Variable	Main effect of time	Main effect of condition	Interaction time*condition
f0 min	$F_{(1, 94)} = 0.112, p = 0.738$	$F_{(1, 94)} = 7.171, p = 0.009$	$F_{(1, 94)} = 0.663, p = 0.417$
f0 max	$F_{(1, 94)} = 0.351, p = 0.555$	$F_{(1, 94)} = 14.073, p < 0.001$	$F_{(1, 94)} = 0.032, p = 0.859$
f0 variability	$F_{(1, 92)} = 4.155, p = 0.044$	$F_{(1, 92)} = 1.784, p = 0.185$	$F_{(1, 92)} = 3.329, p = 0.071$
f0 mean	$F_{(1, 94)} = 0.036, p = 0.849$	$F_{(1, 94)} = 12.643, p < 0.001$	$F_{(1, 94)} = 1.270, p = 0.263$
f0 range	$F_{(1, 88)} = 0.625, p = 0.431$	$F_{(1, 88)} = 1.204, p = 0.276$	$F_{(1, 88)} = 4.838, p = 0.030$

TABLE 3 Summary of the GLMM analyses for the 3 duration variables, in terms of main effects and interactions.

Variable	Main effect of time	Main effect of condition	Interaction time*condition
Number of syllables	$F_{(1, 88)} = 2.171, p = 0.144$	$F_{(1, 88)} = 0.662, p = 0.418$	$F_{(1, 88)} = 0.013, p = 0.911$
Number of silent pauses	$F_{(1, 88)} = 0.303, p = 0.584$	$F_{(1, 88)} = 1.347, p = 0.249$	$F_{(1, 88)} = 0.009, p = 0.926$
Total time of the presentation	$F_{(1, 88)} = 0.927, p = 0.338$	$F_{(1, 88)} = 0.846, p = 0.360$	$F_{(1, 88)} = 0.847, p = 0.360$
Total speaking time	$F_{(1, 88)} = 2.530, p = 0.115$	$F_{(1, 88)} = 0.058, p = 0.811$	$F_{(1, 88)} = 0.731, p = 0.395$
Speech rate	$F_{(1, 94)} = 1.884, p = 0.173$	$F_{(1, 94)} = 4.020, p = 0.048$	$F_{(1, 94)} = 0.918, p = 0.340$
Net syllable rate	$F_{(1, 93)} = 0.743, p = 0.391$	$F_{(1, 93)} = 10.502, p = 0.002$	$F_{(1, 93)} = 5.676, p = 0.019$
ASD	$F_{(1, 93)} = 0.856, p = 0.357$	$F_{(1, 93)} = 7.260, p = 0.008$	$F_{(1, 93)} = 4.472, p = 0.037$

different in triggering SUDS differences in the post-training public speaking task.

Persuasiveness and charisma

The GLMM analysis for persuasiveness showed a main effect of Condition [$F_{(1, 88)} = 7.461, p = 0.008$], which indicated that Non-VR values were higher than VR values ($\beta = 9.869$, SE = 3.613, $p = 0.008$), revealing an imbalance in the values at pre-test across groups in the form of an offset toward generally higher persuasiveness ratings in the Non-VR group as compared to the VR group (both at pre and post-test). The interactions between Time and Condition were not significant, meaning that the training conditions did not have a significantly different effect on the persuasiveness scores at post-training.

Regarding charisma, the GLMM analysis showed a main effect of Condition [$F_{(1, 88)} = 10.625$, $p = 0.002$], which indicated that in general (both at pre- and post-training), Non-VR values were higher than VR values ($\beta = 12.216$, SE = 3.748, $p = 0.002$). The analysis also showed a significant interaction between Time and Condition [$F_{(1, 88)} = 4.245$, $p = 0.042$], which indicated that both at pre-training and post-training the scores for Charisma of the Non-VR group were significantly higher than of the VR group: pre-training ($\beta = 13.821$, $SE = 3.802$, $p < 0.001$), post-training ($\beta = 10.611$, $SE = 3.854$, $p = 0.007$).

Prosodic parameters

F0 domain

Regarding the f0 domain, five GLMMs were applied to our target variables, namely minimum and maximum f0, f0 variability (in terms of the standard deviation), mean f0 and f0 range. Table 2 shows a summary of those GLMM analyses in terms of main effects (Time and Condition), as well as interactions between Time and Condition. Summarizing, a main effect of Time was obtained only for f0 variability, meaning that the post-training values in both groups were higher than the pre-training values. A main effect of Condition was obtained for 3 variables (namely, f0 min, f0 max, and f0 mean), meaning that the participants in the VR group obtained higher f0 values, and larger f0 ranges across both pre- and post-training phases. A significant interaction was obtained for f0 range but no significant *post-hocs* reached significance.

Tempo domain

Regarding the tempo domain, a set of 7 GLMMs were applied to our target variables, namely total number of syllables, total number of silent pauses, total time of the presentation, total speaking time, the speech rate, the net syllable rate and ASD. Table 3 shows a summary of those GLMM analyses in terms of main effects (Time and Condition), as well as interactions

10.3389/fcomm.2022.910952

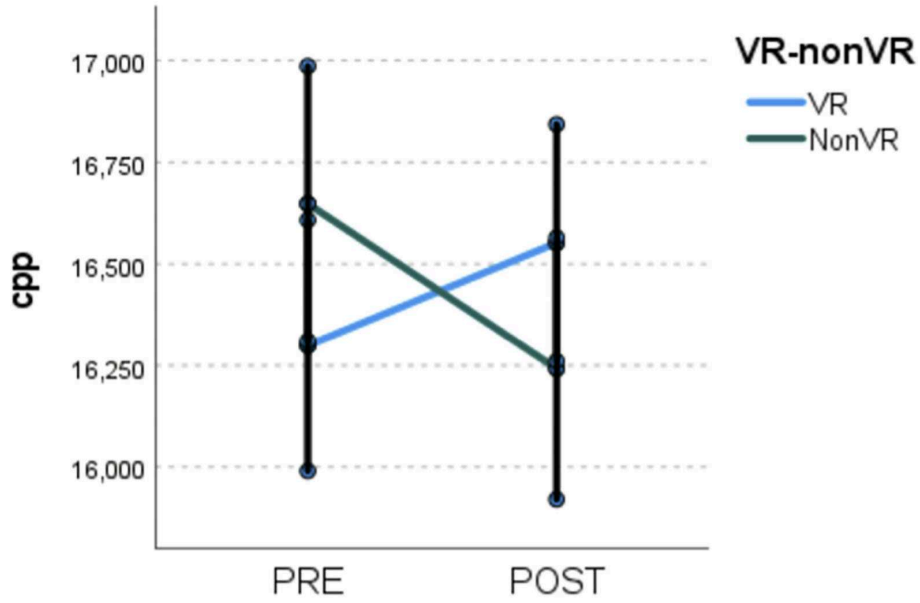


FIGURE 3
Mean CPP values at pre- and post-training, for both VR and Non-VR conditions.

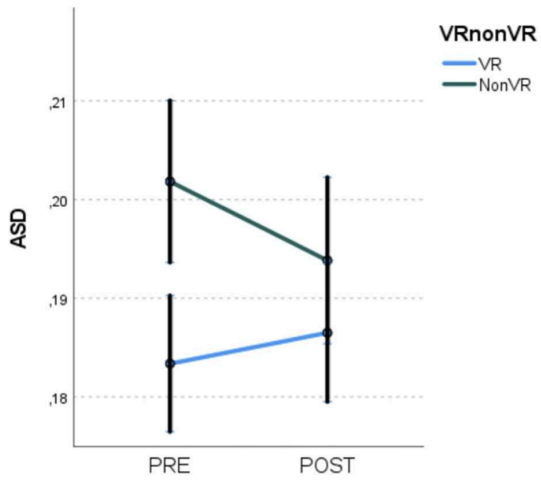


FIGURE 4
Mean ASD values at pre- and post-training, for both VR and Non-VR conditions.

between Time and Condition. Summarizing, no main effects of Time were obtained for any of the parameters of the duration domain. A main effect of Condition was obtained for three variables: speech rate, net syllable rate and ASD, meaning that the participants in the VR group obtained higher speech rate, net syllable rate (or articulation rate) values, and lower ASD values.

The variables that obtained significant interactions were net syllable rate and ASD. For net syllable rate (or articulation rate) in syl/s, the analysis revealed a significant interaction between Time and Condition [$F_{(1, 93)} = 5.676$, $p = 0.019$], which indicated that in the Non-VR group the values were significantly higher at post-training than at pre-training ($\beta = 0.211$, $SE = 0.099$, $p = 0.037$), while no significant differences were found in the VR group ($p = 0.241$). The interaction also showed that at pre-training there was a significant difference between the two groups, showing that the VR group values were higher than the Non-VR group values ($\beta = 0.544$, $SE = 0.143$, $p < 0.001$). With regard to ASD, the GLMM analysis showed a significant interaction between Time and Condition [$F_{(1, 93)} = 4.472$, $p = 0.037$], which indicated that in the Non-VR group the values were significantly lower at post-training than at pre-training ($\beta = 0.008$, $SE = 0.004$, $p = 0.050$), while no significant differences were found in the VR group ($p = 0.358$). VR-group speakers were thus able to maintain their lower ASD levels after training. The interaction also showed that at pre-training there was a significant difference between the two groups, showing that the VR group values were lower than the Non-VR group values ($\beta = 0.018$, $SE = 0.005$, $p = 0.001$). The GLMM analysis also showed a main effect of Condition [$F_{(1, 93)} = 7.260$, $p = 0.008$] which showed that VR values were lower than Non-VR values ($\beta = 0.013$, SE = 0.005, $p = 0.008$). Figures are provided in order to visualize the direction of the effects of the significant interactions. Figures 3, 4 show the mean syllable rate and ASD values obtained in the pre- and post-training tasks across conditions, respectively.

10.3389/fcomm.2022.910952

TABLE 4 Summary of the GLMM analyses for the 10 voice variables, in terms of main effects and interactions.

Variable	Main effect of time	Main effect of condition	Interaction time*condition
h1*-h2*	$F_{(1, 84)} = 0.168, p = 0.683$	$F_{(1, 84)} = 10.774, p = 0.002$	$F_{(1, 84)} = 0.215, p = 0.644$
h1-A3	$F_{(1, 84)} = 18.587, p < 0.001$	$F_{(1, 84)} = 8.523, p = 0.004$	$F_{(1, 84)} = 1.444, p = 0.233$
CPP	$F_{(1, 84)} = 0.334, p = 0.565$	$F_{(1, 84)} = 0.070, p = 0.792$	$F(1,84) = 17.009, p < 0.001$
Harmonicity	$F_{(1, 84)} = 2.197, p = 0.142$	$F_{(1, 84)} = 1.633, p = 0.205$	$F_{(1, 84)} = 0.091, p = 0.763$
CoG	$F_{(1, 84)} = 13.983, p < 0.001$	$F_{(1, 84)} = 0.004, p = 0.951$	$F_{(1, 84)} = 0.443, p = 0.507$
Formant dispersion 1−3	$F_{(1, 84)} = 9.417, p = 0.003$	$F_{(1, 84)} = 1.634, p = 0.205$	$F_{(1, 84)} = 0.023, p = 0.880$
Median pitch	$F_{(1, 82)} = 0.008, p = 0.930$	$F_{(1, 82)} = 14.938, p < 0.001$	$F_{(1, 82)} = 0.514, p = 0.475$
Shimmer	$F_{(1, 84)} = 0.104, p = 0.748$	$F_{(1, 84)} = 3.905, p = 0.051$	$F_{(1, 84)} = 4.195, p = 0.044$
Jitter	$F_{(1, 84)} = 1.845, p = 0.178$	$F_{(1, 84)} = 0.323, p = 0.571$	$F_{(1, 84)} = 3.677, p = 0.059$

Voice quality domain

In the domain of voice quality measurements, a set of 9 GLMMs were applied to our target variables, as explained in section Voice parameters above, namely h1*-h2*, h1-A3, CPP, HNR, CoG, formant dispersion, median pitch, shimmer, and jitter. Table 4 shows a summary of those GLMM analyses in terms of main effects (Time and Condition), as well as interactions between Time and Condition. A set of 9 GLMMs were applied to our target variables, namely h1*-h2*, h1-A3, CPP, HNR, CoG, formant dispersion, median pitch, shimmer, and jitter. Summarizing, a main effect of Time was obtained for 4 variables, namely h1*-h2*, h1-A3, CoG and formant dispersion, meaning that pre-training values were lower at pre-training across groups. A main effect of Condition was obtained for 4 variables, namely h1*-h2*, h1-A3, median pitch and shimmer, meaning that the participants in the VR group obtained higher values compared to the Non-VR group, both at pre and post-trainings.

Significant interactions were obtained for two variables, namely CPP and shimmer and a nearly significant interaction for jitter: For CPP, the GLMM analysis showed a significant interaction between Time and Condition [$F_{(1, 84)} = 17.009$, $p < 0.001$], which indicated that in the Non-VR group the values were significantly lower at post-training than at pre-training ($\beta = 0.351$, $SE = 0.112$, $p = 0.002$), and significantly higher at post-training for the VR group ($p = 0.009$). Regarding shimmer, the GLMM analysis also showed a significant interaction between Time and Condition [$F_{(1, 84)} = 4.195$, $p = 0.044$], which indicated that at pre-test groups were significantly different ($\beta = 0.018$, $SE = 0.008$, $p = 0.039$). The GLMM analysis for jitter showed a near significant interaction between Time and Condition [$F_{(1, 84)} = 3.677$, $p = 0.059$], which indicated that Non-VR values were significantly higher at post-training ($\beta = 0.006$, $SE = 0.003$, $p = 0.035$). Figure 5 shows the mean CPP values obtained in the pre- and post-training tasks across conditions.

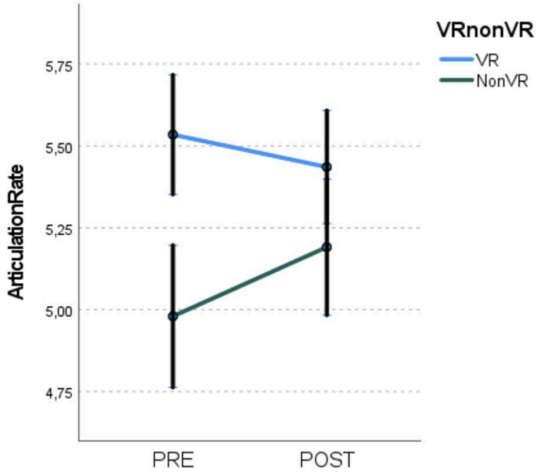


FIGURE 5
Mean articulation rate values at pre- and post-training, for both VR and Non-VR conditions.

Manual gesture rate

The GLMM analysis showed a significant interaction between Time and Condition [$F_{(1, 88)} = 4.796$, $p = 0.031$], but *post-hocs* did not reach significance. No main effects of Time and Condition were found.

Satisfaction questionnaire

Table 5 shows the descriptive results for the 3 questions in the satisfaction questionnaire, separated into VR and Non-VR conditions, on a scale from 1 to 10. As we can see, the responses to the latter two questions yielded higher ratings for the VR group than for the Non-VR group. Specifically, participants of the VR group had on average 0.33 scale points more fun with

TABLE 5 Descriptive results of the satisfaction questionnaire, separated into the VR and Non-VR conditions.

Question	VR group	Non-VR group
Did you feel comfortable participating in the experiment? (1–10)	***Mean*** 8.15	***Mean*** 8.66
	SD 1.22	***SD*** 1.31
	N 19	***N*** 9
Did you have fun? (1–10)	***Mean*** 8.21	***Mean*** 7.88
	SD 0.94	***SD*** 1.46
	N 19	***N*** 9
Do you find the experiment useful for your current oral presentations? (1–10)	***Mean*** 8.10	***Mean*** 6.22
	SD 1.24	***SD*** 2.33
	N 19	***N*** 9

the training task than their Non-VR counterparts and even considered that the perceived usefulness of the VR training was 1.88 scale points higher than their Non-VR counterparts. Yet while the latter difference is statistically significant [$t_{(28)} = 2.891$, $p = 0.004$], the other two are not. We also assessed the reliability of the questionnaire using Cronbach's Alpha. As the number of questions is <10, it is considered that a good reliability score is >= 0.5, and the Cronbach's Alpha score obtained was 0.725.

Discussion

The purpose of this experiment was to examine the impact of a 3-session VR public speaking training on the quality of the oral presentations of a group of 50 secondary school participants when speaking in front of a live audience. Specifically, we assessed the value of two complementary ways of rehearsing speeches, namely rehearsing with a VR audience or rehearsing alone in a room. To achieve this goal, we designed a between-subjects experiment with a pre-training, three training sessions and a post-training so that we could compare pre- to post-training speeches between a VR test condition and a baseline condition of Non-VR training. The duration between pre- and post-training was 5 weeks. One of the key contributions of this study is that it included a comprehensive assessment of the public speaking performance at pre- and post-trainings, specifically by assessing whether presenters in the post-training oral presentation achieved lower levels of anxiety, higher levels of persuasiveness/charisma, and/or a more audience-oriented speech from the point of view of prosodic and gestural features.

First, our results showed that the 3 training sessions reduced the anxiety levels of both VR and Non-VR groups of students to equal degrees in their post-training public speaking task. These results go in line with previous studies where VR trainings proved effective in reducing self-assessed PSA levels of participants, both in clinical settings (e.g., Wallach et al., 2009, 2011; Lister et al., 2010; Lister, 2016; Lindner et al., 2018; Yuen et al., 2019; Zacarin et al., 2019) and in educational settings (e.g., Harris et al., 2002; Heuett and Heuett, 2011; Verano-Tacoronte and Bolívar-Cruz, 2015; Nazligul et al., 2017; Stupar-Rutenfrans et al., 2017). However, this result is not consistent with the hypothesis related to the stronger reduction of self-perceived anxiety in the VR group, as no differences were obtained for the VR and the Non-VR groups. Probably the reason why no differences were found between groups was due to the significant difference at pre-training (a 17-point difference higher for VR) that prevented VR speakers to reduce their self-perceived anxiety to a larger extent.

Second, ratings on persuasiveness and charisma did not result in any significant differences from pre-training to post-training in any of the conditions. This outcome is not consistent with our second hypothesis. As we will discuss later, having obtained no changes in f0 patterns across groups might be the reason behind our results, as greater intonation changes would lead to higher charismatic speech (e.g., (Touati, 1993; Bosker and Kösem, 2017; Niebuhr and Fischer, 2019)), which was not found at post-training for any of the conditions.

Third, with respect to the effects of VR on prosodic parameters, the duration results show that Non-VR speakers significantly raised the articulation rate, i.e., they spoke at a faster pace in the post-training task. A similar change in pace is characteristic of the difference between carefully articulated, and audience-oriented spontaneous speech on the one hand and more self-directed and sloppy read speech on the other (see Jessen, 2007 for the tempo difference between a text-reading exercise and a communicative, spontaneous-speaking task). For ASD, Non-VR participants significantly decreased their values, meaning that they increased their fluency at post-training, but even with this increase were not able to reach the high level of fluency that the VR group was able to maintain at post-training. Voice-quality results show how VR speakers increased their CPP levels from pre-training to post-training speeches. Higher CPP levels are an indication that speakers' voices got clearer and more resonant and confident after training. Importantly, while the VR speakers significantly increased their clarity and resonance, the Non-VR speakers' voices, by contrast, got significantly less clear, resonant, and confident. Very likely this is caused by a reduced vocal effort, i.e., by a softer, less loud voice, produced with lower subglottal pressure.

Thus, overall, our prosody-related results favor an interpretation in which the VR training prevents speakers from falling victim to what Niebuhr and Michalsky (2018) termed the "erosion effect" of repetitive training while, at the same time, it favors a more audience-oriented voice quality in the post-training speeches. The erosion effect caused by repetitive training made the Non-VR speakers' presentations faster and less audience-oriented and their voices less powerful.

This finding is consistent with Niebuhr and Michalsky (2018) who also found that, compared to a control group of speakers who practiced their presentations without VR support, those speakers who could practice with VR support were significantly better able to suppress any negative effects of repetitive rehearsing on their speech prosody—and even improved in some aspects of their speech prosody. Since the lower the *jitter* value the more harmonic, less trembling and creaky the voice is, which suggests that the speakers of the VR group developed at post-training a clearer, stronger and less "shaky" voice, as it was also found by Notaro et al. (2021). Four variables obtained a main effect of Time h1*-h2*, h1-A3, CoG and formant dispersion, meaning that values of both conditions were lower at pre-training. As for a main effect of condition h1*-h2*, h1-A3, and median pitch values were generally higher for the VR condition.

Regarding the duration results, Non-VR speakers significantly raised the articulation rate, i.e., they spoke at a faster pace in the post-training task. For ASD, Non-VR participants significantly decreased their values, meaning that they reduced their fluency at post-training. The Non-VR group thus showed talking faster (>art. rate) and reduce syllable durations (<ASD), probably as a function of rate and fewer pitch accents. All in all, this is in our coaching experience the typical constellation of a bored, uninterested, routine presentation—that does not aim to get a message across to an audience but only to put words into sound.

Surprisingly, our results showed no significant changes across groups on f0 values, meaning that intonation patterns did not change due to VR. At first glance this is inconsistent with the results of Remacle et al. (2021) where teachers performed the same lesson in class and with a virtual audience, or with the results of Niebuhr and Michalsky (2018) where participants had to train persuasive investor pitches with and without a VR audience. The important difference to the present study is, however, that both Niebuhr and Michalsky (2018) and Remacle et al. (2021) analyzed the prosody that speakers showed *during* VR immersion and not after it. As we already highlighted in the Introduction, to our knowledge our experiment is the first to analyze what happens (prosodically) when speakers take off the VR glasses and speak again to a live audience. In fact, as we report in a recent paper on the characteristics of speech during VR public speaking sessions (Valls-Ratés et al., 2021), the prosodic changes that we found when speakers perform public speaking tasks using VR (and Non-VR) are largely consistent with both Niebuhr and Michalsky (2018) and Remacle et al. (2021). F0-related melodic changes can basically be learned through training, as it has been demonstrated by Niebuhr and Neitsch (2020), where the training condition (unlike in our VR condition) included an explicit visualization and color-coded real-time evaluation of speech melody.

Fourth, regarding the use of gesture from pre- to post-training speeches, we did not find significant differences in the post-training task across conditions. We expected to observe a higher rate of gestures as a consequence of the more audience-oriented prosody observed in the VR condition, because research shows that "prominent parts of gestures (or gesture 'hits') tend to align with prosodically prominent parts of speech or pitch accents" (Cravotta et al., 2019, p. 1; see also, Shattuck-Hufnagel et al., 2007; Adrian and Clark, 2011; Loehr, 2012; Esteve-Gibert and Prieto, 2013; Esteve-Gibert et al., 2017). Therefore, our hypothesis regarding an increase in gesture rate for the VR condition is not supported.

Finally, an important result of our investigation is that 17-year-old students found the VR public speaking training (even in its basic, unguided form) more valuable to face their upcoming oral projects than the comparable, traditional rehearsing method without VR. This is also in line with other previous investigations by Kryston et al. (2021), Vallade et al. (2020) and Rodero and Larrea (2022). Thus, promoting more realistic and meaningful ways of individually rehearsing oral skills may enhance the whole experience of delivering a speech with regular and high-quality practice providing a cost-effective practice for education (Merchant et al., 2014; Boetje and van Ginkel, 2021) and increasing students' motivation (Buttussi and Chittaro, 2018; Parong and Mayer, 2018). As we mentioned before, dealing with a high number of students per class and the extensive course curricula makes it extremely difficult for teachers to dedicate hours to enhancing oral skills in-class. Therefore, adopting VR technology could be of great help to make students rehearse individually and encourage them to practice oral skills regularly so as to become more confident and self-aware of their communicative strengths (Merchant et al., 2014; Van Ginkel et al., 2019) and acquire a more charismatic speech (Niebuhr and Michalsky, 2018; Niebuhr and Tegtmeier, 2019) in front of live audiences.

In summary, our study highlights the boosting effects of VR in terms of a handful of duration and voice quality parameters. In general, even though VR leads to preventing the erosion effect and to the use of a more clear and resonant voice after training, we need to acknowledge that this gain in audience-oriented prosody and public speaking confidence that the VR technology achieves, probably based on the presence effect (Slater and Sanchez-Vives, 2003; see section A complementary solution: Empirical evidence on the effects of VR for boosting public speaking skills), was not enough to obtain positive results in many of the other variables that were analyzed within prosodic parameters when the VR-trained speakers were in front of a live audience.

Moreover, a lower SUDS and a more clear voice quality achieved by the VR group were not enough to boost

persuasiveness and charisma scores after the training sessions. Therefore, the match that we expected to see between a more charismatic style in terms of prosodic parameters and the ratings on persuasiveness and charisma was not obtained and we can conclude that the changes in prosodic cues triggered by the VR training were not sufficient to promote a gain in those ratings.

The present study has some limitations. First, the study would have benefitted from a larger sample, which could have yielded more robust results and, thus, a clearer picture of how VR training sessions affect 17-year-old's public-speaking abilities. Second, even though anxiety was controlled through the use of the SUDS scale, a self-assessed measure, adding more objective instruments like electrophysiological measures would allow us to obtain a more fine-grained picture of the anxiety assessment of our participants and compare them with the subjective assessments. Third, in relation to persuasiveness and charisma, raters intuitively assessed the persuasiveness of the message. Even though all speeches contained at least two arguments, we acknowledge that we did not analyze or control for the strength of the arguments nor the rhetorical strategies used by each of the participants (cf. the Charismatic Leadership Tactics of Antonakis et al., 2011), which might have had an influence on the ratings. Fourth, in order to obtain positive effects on charisma and persuasiveness, as well as on f0 parameters, the study could have added more (or longer) training sessions, together with explicit feedback strategies. We believe that giving specific instructions or using feedback strategies to participants (like in Niebuhr and Neitsch, 2020), could change the results at post-training, as seen in other studies (Chollet et al., 2015; Van Ginkel et al., 2019). Future longitudinal studies could be carried out in order to control for the students' perception of enjoyment and usefulness while using VR to ascertain whether the strong value that they assign to VR would remain constant or it is a result of the technology novelty. All in all, designing longer training sessions, longer periods of training, and adding feedback strategies could be regarded as future aims both in research and in practice.

In conclusion, the results of this study serve as a good starting point to continue developing our knowledge about the relationship between VR public speaking practice in secondary school education, self-confidence and the expected improvement in the quality of oral presentations.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Ethics statement

This study was approved by the Commission Project Ethical Revision (Comissió Institucional de Revisió Ètica de Projectes CIREP-UPF) and Recercaixa Project [2017 ACUP 00249] ethical approval. Written informed consent obtained from each participant and/or their legal representative, as appropriate.

Author contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

Funding

This work benefited from funding awarded by the Spanish Ministry of Economy and Competitiveness (PGC2018-097007-B-I00, PID2021-123823NB-I00) and the Generalitat de Catalunya (2017 SGR_971). We also acknowledge the support from the Recercaixa Project (RecerCaixa 2017ACUP 00249) and the Department of Translation at Universitat Pompeu Fabra through a 1-year doctoral grant to ÏV-R.

Acknowledgments

Thank you, young girls and boys of the four high-schools, for believing in the experiment and being so respectful throughout the study. Also to the high-school boards and teachers for being so supportive of the project. Thank you to Florence Baills, Mariia Pronina, and Patrick Rohrer (members of the GrEP group) for your help during data collection. And also GrEP group colleagues Júlia Florit Pons, Xiaotong Xi and Yuan Zhang for your great help with statistics. Thank you, board members, for being so patient and motivated to be part of the experiment. Thanks to Gemma Balaguer Fort, Elisenda Bernal, Gemma Boleda, and Emma Rodero for contributing to our research as committee members of the MA thesis and Ph.D. research defense, your questions have been so valuable. Finally, a special thanks to the 15 raters that assessed the persuasiveness and charisma of all participants.

Conflict of interest

Author ON is CEO of the speech-technology company AllGoodSpeakers ApS. Please see https://oliverniebuhr.com/conflict-of-interest.html for further information.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Adrian, T., and Clark, R. (2011). Audience perceptions of charismatic and non-charismatic oratory: the case of management gurus. *Leadership Q.* 22, 22–32. doi: 10.1016/j.leaqua.2010.12.004

Antonakis, J., Fenley, M., and Liechti, S. (2011). Can charisma be taught? Tests of two interventions. *Acad. Manage. Learn. Educ.* 10, 374–396. doi: 10.5465/amle.2010.0012

Armel, K. C., and Ramachandran, V. S. (2003). Projecting sensations to external objects: evidence from skin conductance response. *Proc. R. Soc. London Ser. B* 270, 1499–1506. doi: 10.1098/rspb.2003.2364

Aymerich-Franch, L., and Bailenson, J. (2014). "The use of doppelgangers in virtual reality to treat public speaking anxiety: a gender comparison," in *Proceedings of the International Society for Presence Research Annual Conference* (Vienna).

Bailey, E. (2018). A historical view of the pedagogy of public speaking. *Voice Speech Rev.* 13, 31–42. doi: 10.1080/23268263.2018.1537218

Banse, R., and Scherer, K. R. (1996). Acoustic profiles in vocal emotion expression. *J. Pers. Soc. Psychol.* 70, 614. doi: 10.1037/0022-3514.70.3.614

Beatty, M. J. (1998). Public speaking apprehension, decision-making errors in the selection of speech introduction strategies and adherence to strategy. *Commun. Educ.* 37, 297–311. doi: 10.1080/03634528809378731

Beatty, M. J., and Behnke, R. R. (1991). Effects of public speaking trait anxiety and intensity of speaking task on heart rate during performance. *Hum. Commun. Res.* 18, 147–176. doi: 10.1111/j.1468-2958.1991.tb00542.x

Beatty, M. J., and Clair, R. P. (1990). "Decision rule orientation and public speaking apprehension," in *Communication, Cognition, and Anxiety*, ed M. Booth-Butterfield (Newbury Park: Sage), 105–116.

Behnke, R. R., and Sawyer, C. R. (1999). Milestones of anticipatory public speaking anxiety, *Commun. Educ.* 48, 165–172. doi: 10.1080/03634529909379164

Boersma, P., and Weenink, D. (2007). *PRAAT: Doing Phonetics by Computer (Version 5.2.34)*. [Computer Software].

Boetje, J., and van Ginkel, S. (2021). The added benefit of an extra practice session in virtual reality on the development of presentation skills: a randomized control trial. *J. Computer Assisted Learn.* 37, 253–264. doi: 10.1111/jcal.12484

Boettcher, J., Carlbring, P., Renneberg, B., and Berger, T. (2013). Internet-based interventions for social anxiety disorder - an overview. *Verhaltenstherapie* 23, 160–168. doi: 10.1159/000354747

Bosker, H. R., and Kösem, A. (2017). "An entrained rhythm's frequency, not phase, influences temporal sampling of speech," in *Proceedings of Interspeech 2017* (Stockholm), 2416–2420. doi: 10.21437/Interspeech.2017-73

Bratman, G., Daily, G., Levy, B., and Gross, J. (2015). The benefits of nature experience: Improved affect and cognition. *Landsc. Urban Plan.* 138, 41–50. doi: 10.1073/pnas.1510459112

Brown, T., and Morrissey, L. (2004). The effectiveness of verbal self-guidance as a transfer of training intervention: its impact on presentation performance, self-efficacy and anxiety. *Innov. Educ. Teach. Int.* 41, 255–271. doi: 10.1080/14703290410001733302

Buttussi, F., and Chittaro, L. (2018). Effects of different types of virtual reality display on presence and learning in a safety training scenario. *IEEE Trans. Vis. Comput. Graph.* 24, 1063–1076. doi: 10.1109/TVCG.2017.2653117

Choi, C. W., Honeycutt, J. M., and Bodie, G. D. (2015). Effects of imagined interactions and rehearsal on speaking performance. *Commun. Educ.* 64, 25–44. doi: 10.1080/03634523.2014.978795

Chollet, M., Wörtwein, T., Morency, L. P., Shapiro, A., and Scherer, S. (2015). "Exploring feedback strategies to improve public speaking: an interactive virtual audience framework," in *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (ACM) (Osaka), 1143–1154.

Cravotta, A., Busa, M. G., and Prieto, P. (2019). Effects of encouraging the use of gestures on speech. *J. Speech Lang. Hearing Res.* 62, 1–16. doi: 10.1044/2019_JSLHR-S-18-0493

Dalgarno, B., and Lee, M. (2010). What are the learning affordances of 3-D virtual environments?. *Br. J. Educ. Technol.* 41, 10–32. doi: 10.1111/j.1467-8535.2009.01038.x

Daly, J. A., Vangelisti, A. L., and Weber, D. J. (1995). Speech anxiety affects how people prepare speeches: a protocol analysis of the preparation processes of speakers. *Commun. Monogr.* 62, 383–397. doi: 10.1080/03637759509376368

Daniels, M. M. (2021). "Usability assessment of virtual reality as a training tool for oral presentation," in *IOP Conference Series: Materials Science and Engineering* (Sanya).

Daniels, M. M., Palaoag, T., and Daniels, M. (2020). "Efficacy of virtual reality in reducing fear of public speaking: a systematic review," in *International Conference on Information Technology and Digital Applications* (Sanya), 803.

De Jong, N. H., and Wempe, T. (2009). Praat script to detect syllable nuclei and measure speech rate automatically. *Behav. Res. Methods* 41, 385–390. doi: 10.3758/BRM.41.2.385

Ekman, P., and Friesen, W. V. (1969). The repertoire or nonverbal behavior: categories, origins, usage and coding. *Semiotica*. 1, 49–98. doi: 10.1515/semi.1969.1.1.49

Esteve-Gibert, N., Borràs-Comes, J., Asor, E., Swerts, M., and Prieto, P. (2017). The timing of head movements: the role of prosodic heads and edges. *J. Acoust. Soc. Am.* 141, 4727. doi: 10.1121/1.4986649

Esteve-Gibert, N., and Prieto, P. (2013). Prosody signals the emergence of intentional communication in the first year of life: evidence from Catalan-babbling infants. *J. Child Lang.* 40, 919–944. doi: 10.1017/S0305000912000359

Fox Cabane, O. (2013). *The Charisma Myth - How Anyone Can Master the Art and Science of Personal Magnetism*. New York, NY: Penguin.

Frisby, B. N., Kaufmann, R., Vallade, J. I., Frey, T. K., and Martin, J. C. (2020). Using virtual reality for speech rehearsals: an innovative instructor approach to enhance student public speaking efficacy. *Basic Commun. Course Ann.* 32, Article 6. Available online at: https://ecommons.udayton.edu/cgi/viewcontent.cgi?article=1587&context=bcca

Gao, D. (2022). Oral english training based on virtual reality technology. *Eng. Intelligent Syst.* 30, 49–54. Available online at: https://website-eis.crlpublishing.com/index.php/eis/article/view/1647

Goberman, A. M., Hughes, S., and Haydock, T. (2011). Acoustic characteristics of public speaking: anxiety and practice effects. *Speech Commun.* 53, 867–876. doi: 10.1016/j.specom.2011.02.005

Gross, J. J., and John, O. P. (2003). Individual differences in two emotion regulation processes: Implications for affect, relationships, and well-being. *J. Pers. Soc. Psychol.* 85, 348–362. doi: 10.1037/0022-3514.85.2.348

Gruber, A., and Kaplan-Rakowski, R. (2020). "User experience of public speaking practice in virtual reality," in *Cognitive and Affective Perspectives on Immersive Technology in Education*, ed R. Zheng (Hershey: IGI Global), 235–249.

Harris, S. R., Kemmerling, R. L., and North, M. (2002). Brief virtual reality therapy for public speaking anxiety. *Cyberpsychol. Behav.* 5, 543–550. doi: 10.1089/109493102321018187

Heuett, B. L., and Heuett, K. B. (2011). Virtual reality therapy: a means of reducing public speaking anxiety. *Int. J. Humanities Soc. Sci.* 1, 1–6. doi: 10.3390/jpm10010014

Howard, M. C., and Gutworth, M. B. (2020). A meta-analysis of virtual reality training programs for social skill development. *Comput. Educ.* 144, 103707. doi: 10.1016/j.compedu.2019.103707

Jessen, M. (2007). Forensic reference data on articulation rate in German. *Sci. Justice J. Forensic Sci. Soc.* 47, 50–67. doi: 10.1016/j.scijus.2007.03.003

Kahlon, S., Lindner, P., and Nordgreen, T. (2019). Virtual reality exposure therapy for adolescents with fear of public speaking: a non-randomized feasibility and pilot study. *Child Adolesc. Psychiatry Ment. Health* 13, 47. doi: 10.1186/s13034-019-0307-y

Kendon, A. (2004) *Gesture: Visible Action as Utterance*. Cambridge: Cambridge University Press.

Koo, T. K., and Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J. Chiropr. Med.* 15, 155–163. doi: 10.1016/j.jcm.2016.02.012

Kryston, K., Goble, H., and Eden, A. (2021). Incorporating virtual reality training in an introductory public speaking course. *J. Commun. Pedagogy* 4, 133–151. doi: 10.31446/JCP.2021.1.13

Lane, H. C., Hays, M. J., Core, M. G., and Auerbach, D. (2013). Learning intercultural communication skills with virtual humans: feedback and fidelity. *J. Educ. Psychol.* 105, 1026–1035. doi: 10.1037/a0031506

LeFebvre, L. E., LeFebvre, L., and Allen, M. (2020). "Imagine all the people": imagined interactions in virtual reality when public speaking. *Imagin. Cogn. Pers.* 40, 189–222. doi: 10.1177/0276236620938310

Legault, J., and Zhao, J., and Chi, Y., and Chen, W., and Klippel, A., and Li, P. (2019). Immersive virtual reality as an effective tool for second language vocabulary learning. *Languages* 4, 13. doi: 10.3390/languages4010013

Lindner, P., Dagöö, J., Hamilton, W., Miloff, A., Andersson, G., Schill, A., et al. (2020). Virtual reality exposure therapy for public speaking anxiety in routine care: a single-subject effectiveness trial. *Cogn. Behav. Ther.* 50, 67–87. doi: 10.1080/16506073.2020.1795240

Lindner, P., Miloff, A., Fagernäs, S., Andersen, J., Sigeman, M., Andersson, G., et al. (2018). Therapist-led and self-led one-session virtual reality exposure therapy for public speaking anxiety with consumer hardware and software: a randomized controlled trial. *J. Anxiety Disord.* 61, 45–54. doi: 10.1016/j.janxdis.2018.07.003

Lister, H. (2016). *The effect of virtual reality exposure on fear of public speaking using cloud-based software* (Thesis), University of New Brunswick, Fredericton, NB, Canada.

Lister, H. A., Piercey, C. D., and Joordens, C. (2010). The effectiveness of 3-D video virtual reality for the treatment of fear of public speaking. *J. Cyber Ther. Rehabil.* 3, 375–381. Available online at: https://www.scopus.com/inward/record.uri?eid=2-s2.0-79960359620&partnerID=40&md5=a5fe8e42ae54ec8f46e1272ef96a3748

Liu, X., and Xu, Y. (2014). Body size projection and its relation to emotional speech-evidence from Mandarin Chinese. *Proc. Speech Prosody* 2014, 974–977. doi: 10.21437/SpeechProsody.2014-184

Loehr, D. (2012). Temporal, structural, and pragmatic synchrony between intonation and gesture. *Lab. Phonol.* 3, 71–89. doi: 10.1515/lp-2012-0006

Lövgren, T., and Doorn, J. V. (2005). "Influence of manipulation of short silent pause duration on speech fluency," in *Proceedings of Disfluency in Spontaneous Speech Conference* (Aix-en-Provence), 123–126.

Macinnis, C., Mackinnon, S., and Macintyre, P. (2010). The illusion of transparency and normative beliefs about anxiety during public speaking. *Curr. Res. Soc. Psychol.* 15, 42–52. Available online at: http://www.uiowa.edu/~grpproc/crisp/crisp.html

McCroskey, J. (1982). Oral communication apprehension: a reconceptualization. *Ann. Int. Commun. Assoc.* 6, 136–170. doi: 10.1080/23808985.1982.11678497

McNeill, D. (1992). *Hand and Mind: What Gestures Reveal About Thought*. University of Chicago Press.

Menzel, K. E., and Carrell, L. J. (1994). The relationship between preparation and performance in public speaking. *Commun. Educ.* 43, 17–26. doi: 10.1080/03634529409378958

Merchant, Z., Goetz, E., Cifuentes, L., Keeney-Kennicutt, W., and Davis, T. (2014). Effectiveness of virtual reality-based instruction on students' learning outcomes in K-12 and higher education: a meta-analysis. *Comput. Educ.* 70, 29–40. doi: 10.1016/j.compedu.2013.07.033

Mikropoulos, T., and Natsis, A. (2011). Educational virtual environments: a ten-year review of empirical research (1999-2009). *Comput. Educ.* 56, 769–780. doi: 10.1016/j.compedu.2010.10.020

Morreale, S. P., Osborn, M. M., and Pearson, J. C. (2000). Why communication is important: a rationale for the centrality of the study of communication. *J. Assoc. Commun. Administr.* 29, 125. Available online at: https://www.academia.edu/29493673/Why_Communication_is_Important_A_Rationale_for_the_Centrality_of_the_Study_of_Communication

Nazligul, M., Yilmaz, M., Gulec, U., Gözcü, M., O'Connor, R., and Clarke, P. (2017). "Overcoming public speaking anxiety of software engineers using virtual reality exposure therapy," in *24th European Conference on Systems, Software and Services Process Improvement (EuroSPI 2017)* (Ostrava), 191–202.

Niebuhr, O. (2021). "Advancing higher-education practice by analyzing and training students' vocal charisma: evidence from a Danish field study," in *7th International Conference on Higher Education Advances (HEAd'21)* (Valencia), 743–751.

Niebuhr, O., Brem, A., Michalsky, J., and Neitsch, J. (2020). What makes business speakers sound charismatic? A contrastive acoustic-melodic analysis of Steve Jobs and Mark Zuckerberg. *Cadernos de Linguística* 1, 1–40. doi: 10.25189/2675-4916.2020.v1.n1.id272

Niebuhr, O., and Fischer, K. (2019). "Do not hesitate! - unless you do it shortly or nasally: how the phonetics of filled pauses determine their subjective frequency and perceived speaker performance," in *Proceedings of Interspeech 2019* (Graz), 544–548. doi: 10.21437/Interspeech.2019-1194

Niebuhr, O., and Michalsky, J. (2018). "Virtual reality simulations as a new tool for practicing presentations and refining public-speaking skills," in *Proceedings of the 9th International Conference on Speech Prosody 2018* [International Speech Communication Association (ISCA)] (Poznan), 309–313.

Niebuhr, O., and Neitsch, J. (2020). "Digital rhetoric 2.0: how to train charismatic speaking with speech-melody visualization software," in *Lecture Notes in Computer Science*, Vol. 12335, *Speech & Computer*, eds A. Karpov and R. Potapova (New York, NY: Springer Nature), 357–368.

Niebuhr, O., and Tegtmeier, S. (2019). "Virtual reality as a digital learning tool in entrepreneurship: how virtual environments help entrepreneurs give more charismatic investor pitches, FGF studies in small business and entrepreneurship," in *Digital Entrepreneurship*, eds R. Baierl, J. Behrens, and A. Brem (Cham: Springer), 123–158.

North, M. M., North, S. M., and Coble, J. R. (2015). Virtual reality therapy: an effective treatment for the fear of public speaking. *Int. J. Virtual Real.* 3, 1–6. doi: 10.20870/IJVR.1998.3.3.2625

Notaro, A., Capraro, F., Pesavento, M., Milani, S., and Busà, M. G. (2021). "Effectiveness of VR immersive applications for public speaking enhancement," in *Proceedings of IS&T International Symposium on Electronic Imaging: Image Quality and System Performance XVIII* (Springfield, MA), 2021, 294-1–294-7.

Parong, J., and Mayer, R. E. (2018). Learning science in immersive virtual reality. *J. Educ. Psychol.* 110, 785–797. doi: 10.1037/edu0000241

Pearson, J. C., Child, J. T., and Kahl, D. H. Jr. (2006). Preparation meeting opportunity: how do college students prepare for public speeches? *Commun. Q.* 54, 351–366. doi: 10.1080/01463370600878321

Peeters, D. (2019). Virtual reality: a game-changing method for the language sciences. *Psychon. Bull. Rev.* 26, 894–900. doi: 10.3758/s13423-019-01571-3

Rasipuram, S., Rao, S. P., and Jayagopi, D. B. (2016). "Automatic prediction of fluency in interface-based interviews," in *2016 IEEE Annual India Conference (INDICON)* (Bangalore: IEEE), 1–6.

Remacle, A., Bouchard, S., Etienne, A., Rivard, M., and Morsomme, D. (2021). A virtual classroom can elicit teachers' speech characteristics: evidence from acoustic measurements during *in vivo* and *in virtuo* lessons, compared to a free speech control situation. *Virtual Real.* 25, 935–944. doi: 10.1007/s10055-020-00491-1

Rocklage, M. D., Rucker, D. D., and Nordgren, L. F. (2018). Persuasion, emotion, and language: the intent to persuade transforms language via emotionality. *Psychol. Sci.* 29, 749–760. doi: 10.1177/0956797617744797

Rodero, E., and Larrea, O. (2022). Virtual reality with distractors to overcome public speaking anxiety in university students. [Realidad virtual con distractores para superar el miedo a hablar en público en universitarios]. *Comunicar* 72, 87–99. doi: 10.3916/C72-2022-07

Rogers, R., Wooley, J., Sherrick, B., David Bowman, N., and Oliver, M. B. (2017). Fun versus meaningful video game experiences: a qualitative analysis of user responses. *Comput. Game J.* 6, 63–79. doi: 10.1007/s40869-016-0029-9

Rubin, R. B., Rubin, A. M., and Jordan, F. F. (1997). Effects of instruction on communication apprehension and communication competence. *Commun. Educ.* 46, 104–114. doi: 10.1080/03634529709379080

Ruscella, J. J. (2019). *Virtual Reality and Job Interviews*. Academic Minute. Available online at: https://academicminute.org

Safir, M. P., Wallach, H. S., and Bar-Zvi, M. (2012). Virtual reality cognitive-behavior therapy for public speaking anxiety: one-year follow-up. *Behav. Modif.* 36, 235–246. doi: 10.1177/0145445511429999

Sakib, M. N., Chaspari, T., and Behzadan, A. (2019). "Coupling virtual reality and physiological markers to improve public speaking performance," in *19th International Conference on Construction Applications of Virtual Reality (CONVR2019)* (Bangkok), 171–180.

Shattuck-Hufnagel, S., Yasinnik, Y., Veilleux, N., and Renwick, M. (2007). "A method for studying the time alignment of gestures and prosody in American English: 'Hits' and pitch accents in academic-lecture-style speech," in *Fundamentals of Verbal and Nonverbal Communication and the Biometric Issue*, Vol. 18, eds A. Esposito, M. Bratanic, E. Keller, and M. Marinaro (Amsterdam: IOS Press), 34–44.

Slater, M., and Sanchez-Vives, M. V. (2003). Enhancing our lives with immersive virtual reality. *Front. Robot. AI* 3, 74. doi: 10.3389/frobt.2016.00074

Smith, T. E., and Frymier, A. B. (2006). Get'real': does practicing speeches before an audience improve performance?. *Commun. Q.* 54, 111–125. doi: 10.1080/01463370500270538

Spring, R., Kato, F., and Mori, C. (2019). Factors associated with improvement in oral fluency when using video-synchronous mediated communication with native speakers. *Foreign Lang. Ann.* 52, 87–100. doi: 10.1111/flan.12381

Stupar-Rutenfrans, S., Ketelaars, L., and Van Gisbergen, M. S. (2017). Beat the fear of public speaking: mobile 360° video virtual reality exposure training in home environment reduces public speaking anxiety. *Cyberpsychol. Behav. Soc. Netw.* 20, 624–633. doi: 10.1089/cyber.2017.0174

Takac, M., Collett, J., Blom, K. J., Conduit, R., Rehm, I., and De Foe, A. (2019). Public speaking anxiety decreases within repeated virtual reality training sessions. *PLoS ONE* 14, e0216288. doi: 10.1371/journal.pone.0216288

Thrasher, T. (2022). The impact of virtual reality on L2 French learners' language anxiety and oral comprehensibility: an exploratory study. *CALICO J.* 39, 219–238. doi: 10.1558/cj.42198

Touati, P. (1993) "Prosodic aspects of Political rhetoric," in *Proc. ESCA Workshop on Prosody* (Lund), 168–171.

Tse, A. Y. (2012). Glossophobia in university students of Malaysia. *Int. J. Asian Soc. Sci.* 2, 2061–2073. Available online at: https://archive.aessweb.com/index.php/5007/article/view/2374

Vallade, J. I., Kaufmann, R., Frisby, B. N., and Martin, J. C. (2020). Technology acceptance model: investigating students' intentions toward adoption of 360° videos for public speaking rehearsals. *Commun. Educ.* 70, 127–145. doi: 10.1080/03634523.2020.1791351

Valls-Ratés, I., Niebuhr, O., and Prieto, P. (2021). "Effects of public speaking virtual reality trainings on prosodic and gestural features," in *Proceedings of 1st International Conference on Tone and Intonation (TAI)* (Sønderborg), 214–218. doi: 10.21437/SpeechProsody.2022-33

Van Ginkel, S., Gulikers, J., Biemans, H., Noroozi, O., Roozen, M., Bos, T., et al. (2019). Fostering oral presentation competence through a virtual reality-based task for delivering feedback. *Comput. Educ.* 134, 78–97. doi: 10.1016/j.compedu.2019.02.006

Van Ginkel, S., Ruiz, D., Mononen, A., Karaman, A. C., de Keijzer, A., and Sitthiworachart, J. (2020). The impact of computer-mediated immediate feedback on developing oral presentation skills: an exploratory study in virtual reality. *J. Computer Assisted Learn.* 36, 412–422. doi: 10.1111/jcal.12424

Verano-Tacoronte, D., and Bolívar-Cruz, A. (2015). "La confianza para hablar en público entre los estudiantes universitarios," in *Conference: XXIX AEDEM Annual Meeting, June 2015* (Donostia).

Wallach, H. S., Safir, M. P., and Bar-Zvi, M. (2009). Virtual reality cognitive behavior therapy for public speaking anxiety: a randomized clinical trial. *Behav. Modif.* 33, 314–338. doi: 10.1177/0145445509331926

Wallach, H. S., Safir, M. P., and Bar-Zvi, M. (2011). Virtual reality exposure versus cognitive restructuring for treatment of public speaking anxiety: a pilot study. *ISR. J. Psychiatry Relat. Sci.* 48, 91–97.

Wilsdon, L., and Fullwood, C. (2017). The effect of immersion and presence in a virtual reality public speaking task. *Annu. Rev. CyberTherapy Telemed.* 17, 211–213. Available online at: http://hdl.handle.net/2436/622249

Wolpe, J. (1969). *The Practice of Behavior Therapy, Pergamon General Psychology Series, 1*. 4th Edn. Elmsford, NY: Pergamon Press.

Wolpe, J. (1990). *Pergamon General Psychology Series, 1. The Practice of Behavior Therapy, 4th Edn*. Elmsford, NY, US: Pergamon Press.

Xu, Y. (2013). "ProsodyPro - a tool for large-scale systematic prosody analysis," in *Proceedings of Tools and Resources for the Analysis of Speech Prosody (TRASP 2013)* (Aix-en-Provence), 7–10.

Xu, Y., Park, H., and Baek, Y. (2011). A new approach toward digital storytelling: an activity focused on writing selfefficacy in a virtual learning environment. *Educ. Technol. Soc.* 14, 181–191. Available online at: http://www.ifets.info/journals/14_4/16.pdf

Yuen, E. K., Goetter, E. M., Stasio, M. J., Ash, P., Mansour, B., McNally, E., et al. (2019). A pilot of acceptance and commitment therapy for public speaking anxiety delivered with group videoconferencing and virtual reality exposure. *J. Contextual Behav. Sci.* 12, 47–54. doi: 10.1016/j.jcbs.2019.01.006

Zacarin, M. R. J., Borloti, E., and Haydu, V. B. (2019). Behavioral therapy and virtual reality exposure for public speaking anxiety. *Trends Psychol.* 27, 491–507. doi: 10.9788/TP2019.2-14

Zellin, M., Mühlenen, A., and von Müller, H. J. (2014). Long-term adaptation to change in implicit contextual learning. *Psychon. Bull. Rev.* 21, 1073–1079. doi: 10.3758/s13423-013-0568-z

Appendix

Instructions for the PRE- and POST-training public speaking tasks

ENGLISH VERSION

PREPARATION TIME: **2 MINUTES**

SPEECH DURATION: **2 MINUTES**

Situation: Three representatives of the Education Department have come to your high-school to listen to the proposals of a group of students. They're thinking of assigning more budget to school trips.

Your claim is that adolescents need to **spend more time in nature** and not so many hours inside the city schools.

In order to argue in favor of your proposal, you have prepared a list of studies with data that will allow you to **convince the representatives** to assign more budget to this field.

- More than 50% of the population lives nowadays in urban areas. It is estimated that in 2050 the number will increase up to a 70% (Bratman et al., 2015).
- People surrounded of less trees suffer more stress and higher mortality rates.
- Be surrounded of nature reduces the stress hormone, blood pressure and sugar in the blood.
- Be surrounded of nature increases cardiovascular and metabolic health, concentration and memory.
- Strolling in the forest increases creativity, vitality and relaxation (Finnish Forest Research Institute).

GOOD LUCK!

Instructions for the TRAINING 1 public speakingn task

ENGLISH VERSION

THE HOUSE OF MY DREAMS

PREPARATION TIME: **2 MINUTES**

SPEECH DURATION: **2 MINUTES**

Script that can help you prepare the structure and content of the message:

- Description of the house
- Place
- Why would it be like that?
- What would be essential to be part of the house?
- What would you do in such a house?
- Would you live alone or would you like to share it with other people?

GOOD LUCK!

Instructions for the TRAINING 2 public speaking task

ENGLISH VERSION

IS GRAFFITI ART?

PREPARATION TIME: **2 MINUTES**

SPEECH DURATION: **2 MINUTES**

Script that can help you prepare the structure and content of the message:

- Description of what is a graffiti
- Where do we usually find them
- Who makes them
- Why are they important/necessary or the opposite
- What makes you state that it is art or not and why
- Use examples and personal experience

GOOD LUCK!

Instructions for the TRAINING 3 public speaking task

ENGLISH VERSION

MONEY CAN NOT BUY HAPPINESS

PREPARATION TIME: **2 MINUTES**

SPEECH DURATION: **2 MINUTES**

Script that can help you prepare the structure and content of the message:

- How would you describe happiness?
- What does money buy and what doesn't?
- Richness/Poverty
- What makes you state or negate the topic sentence
- Use examples or experience that can illustrate feelings of happiness

GOOD LUCK!

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 05 October 2022
DOI 10.3389/fpsyg.2022.938916

Check for updates

OPEN ACCESS

EDITED BY
Francesca D'Errico,
University of Bari Aldo Moro, Italy

REVIEWED BY
Yaodong Gu,
Ningbo University, China
Maxwell Barnish,
University of Exeter, United Kingdom
Paul Haydn Morris,
University of Portsmouth,
United Kingdom

*CORRESPONDENCE
Radomír Masaryk
masaryk1@uniba.sk

SPECIALTY SECTION
This article was submitted to Gender, Sex and Sexualities, a section of the journal Frontiers in Psychology

RECEIVED 08 May 2022
ACCEPTED 12 July 2022
PUBLISHED 05 October 2022

CITATION
Masaryk R, Synak N and Belišová M (2022) Is wearing high heels a female mating strategy? Revisiting the original study using qualitative methods. *Front. Psychol.* 13:938916. doi: 10.3389/fpsyg.2022.938916

COPYRIGHT
© 2022 Masaryk, Synak and Belišová. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Is wearing high heels a female mating strategy? Revisiting the original study using qualitative methods

Radomír Masaryk*, Nikola Synak and Michaela Belišová

Institute of Applied Psychology, Faculty of Social and Economic Sciences, Comenius University Bratislava, Bratislava, Slovakia

The 2020 study entitled 'Wearing high heels as female mating strategy' by Pavol Prokop and Jana Švancárová claimed that when females imagined an interaction with an attractive male, their preference for high heels steeply increased, compared with an imagined interaction with an unattractive male. The authors concluded that wearing high heels seem to be a form of sexual signaling by females in intersexual interactions. The present paper revisits this study through a psychological standpoint, rather than a biological one. In addition to proposing hypothetical dating scenarios, as in the original study, we also asked participants about how they went about getting ready to go on a date, the significance of dating to them, and their thinking behind choosing particular outfits for a date. We conducted ten focus groups (N = 50), recruiting from a similar sample of participants to those in the original study. For our study we followed principles of Thematic Analysis to identify the key themes in the narratives related to dating and beautification. We also used the photo elicitation methodology to observe what footwear our participants own. Our data interpretation from these two sources suggests that young women tend to see dates as social events not necessarily leading to sex; that they do not regard high heel shoes as a means of beautification; and that they take account of practical considerations when getting dressed up for a date. Moreover, young women tend to use beautification with caution. We conclude that the relationship between the tendency to use beautification and attractiveness of a potential partner is far from straightforward; and relying on binary responses to hypothetical scenarios does not provide convincing evidence.

KEYWORDS

physical attractiveness, dating, beautification, partner selection, sexual communication

Introduction

In their paper, Prokop and Švancárová (2020) claim that wearing high heels operates as a form of sexual signaling by females in intersexual interactions. They observed that when female participants imagined an interaction with an attractive male, their preference for high heels steeply increased. However, the study design is based on participants imagining having an interaction with one of two different photographs (attractive/unattractive mele) and choosing one of two kinds of footwear (as shown on a photographs (high heels and low heels). Our study was designed to open up alternative explanations, based on psychological, rather than biological approaches. We used qualitative methods to explore how young women view dating events, with focusing in particular on themes associated with the phenomenon of beautification. Our intention was to explore the narratives of similar young women (N = 21 years, SD = 4.67 in the original study) that relate to dating and beautification. Our aim was to gain deeper insights beyond the simple forced choice response used in the original study.

The theoretical foundation of the original paper (Prokop and Švancárová, 2020) is based on the premise that *"human females invest in offspring more than males"* (supported by references to Eibl-Eibesfeldt, 1989; Hewlett, 1992; Geary, 2000) and *"human males, on the other hand, spend more time caring for their offspring than males of virtually all other mammals"* (Bjorklund and Shackelford, 1999; Puts, 2010). According to these sources, parental investment constitutes a basis for female competition over a potential sexual partner. One of the ways for females to achieve this according to Prokop and Švancárová (2020) is beautification, as exemplified by wearing high heels. The parental investment theory (Trivers, 1972) has however, been repeatedly criticized by psychologists on several grounds. Buss (1994) claims that people show a remarkable measure of creativity when it comes to their reproduction strategies. One example is the model of a single long-term relationship combined with several concurring extra-marital affairs. His study involving 10,047 respondents from 37 cultures (Buss, 1989, 1994) provides evidence that women choose their potential partner based on the potential access of men to resources and their willingness to share these resources, as well as their capability to physically protect the family and aspects such as compatibility of their objectives and values. Most importantly, the authors demonstrate that sexual strategies in humans may be complex and multi-faceted.

The association between wearing high heels and sexuality has nevertheless been well documented by several studies. For instance, Morris et al. (2013) found that participants viewed women wearing high heel shoes on point-light videos as significantly more attractive than those with flat shoes. The authors attribute this to increased femininity of gait (reduced stride length and increased rotation and tilt of the hips) and conclude that a woman walking in high heels is a supernormal stimulus. Similar findings were reported in a more recent study by Wade et al. (2022): silhouettes showing women wearing high heels were rated as more attractive, both physically and sexually than the silhouettes of women in flat shoes. The high heel wearers were also seen as more feminine and as of higher status. The 2016 systematic review by Barnish et al. (2018) also confirmed the association between wearing high heels and increased attractiveness and/or an impact on men's behavior. However, they also confirm the well-documented association between wearing high heel wear and negative musculoskeletal health effects. Wade et al. (2022) suggest that the pain and damage is a trade-off for an increase in perceived attractiveness. High heels were also found to be heavily represented in pornographic imagery (Dietz and Evans, 1982), which may suggest that they have become a cultural symbol of sexuality.

Target groups of the studies in question typically focus on female college undergraduate students. Although the college environment is often stereotyped as highly sexualized, the American College Health Association (2018) reported that 32.3% of female college students have never engaged in oral sex and 33.9% of female college students have never engaged in vaginal sex. Hills (2015) ironically states that *"college campuses are portrayed as being this hotbed of hookup culture"*; the author however, claims that it is not necessarily always the case and a significant proportion of college students are much less sexually active than portrayed in popular culture and media. This leads to a question whether reproductive success or sexual pleasure are truly the driving forces behind clothing and footwear decisions for this target group.

Although the relationship between wearing high heels and physical attractiveness has been repeatedly documented, little is known about the wider context of decisions made by those who decide to wear them. Studies of this kind typically used videos, photographs, silhouettes, or point-light videos in order to reduce the situation to a binary choice of wearing high heels or flats. The study by Prokop and Švancárová (2020) used a hypothetical scenario. This raises several methodological issues. For example, one person looked submissive, introverted, shy, and visibly shorter, while the other seemed much more dominant, extroverted, self-confident, and taller. Forced choice between the two images may yield an answer in keeping with social stereotypes but may not indicate real preference. When McDaniel (2005) asked if there is support for the popular myth that young women prefer jerks as dating partners rather than nice guys, she concluded that young women choose to date nice guys when they are perceived as possessing a combination of attractive personality traits. Physical attractiveness is only one of the factors involved in addition to other qualities such as being funny/witty, romantic, exciting, and someone whom their friends might like. Based on this we believe young women in a hypothetical scenario based their decision on imagined qualities and personality traits of young men in the images rather than on perceived physical attractiveness. Moreover, as one of

the reviewers of the present paper suggested in anonymous peer-review, the hypothetical of what would a woman wear to a date with someone they find unattractive is very questionable.

The authors of the original paper suggest that when choosing footwear, a shorter body height is one of the factors (Prokop and Švancárová, 2020). We agree that this could be one of many different factors that come into play when young women choose their outfit for a date. Such decisions are complex and should not be reduced to potential sexual interest. While typical studies in the field attempt to clear the choice of its context, our intention was exactly the opposite: try to examine the choice of footwear within the wider context of dating decisions. The objective of the present paper was to look deeper into narratives expressed by young women both generally as well as specifically when presented with the choice offered in the original paper by Prokop and Švancárová (2020).

Materials and methods

Research sample

The research sample consisted of 50 women aged 19 to 24 years (M = 21.04; SD = 3.25), which was similar to the sample in the original study (M = 21 years, SD = 4.67). Our sample was recruited among students at five universities in Slovakia and one in Czechia. Participants were recruited using university social networks and awarded a small extra credit for coursework. Most of the participants studied humanities, especially psychology (94%). Out of the total number of participants, 34 (68%) reported being in a relationship. The mean number of previous sexual partners ranged from 0 to 13 (M = 2.90; SD = 2.47). The mean height of our participants was 168.61 cm or 5.53 ft (SD = 40.29).

Data collection

Having received ethical approval from the local Institutional Review Board we conducted 10 focus groups consisting of 4 to 6 participants each. Participants who agreed to participate in the study were given a chance to choose from possible time slots to their schedule until we reached the maximum number of 6 for each focus group. Groups were facilitated by the two women researchers who co-authored the present study, to avoid embarrassment while discussing sensitive issues.

In advance of every group, we sent our participants registration forms requesting their contact information, education level, and partnership status. We also asked them to provide a photograph of their entire shoe collection available at the current place of residence (solicited photography). The pictures were used during focus groups to aid memory recall. After each group we sent them a questionnaire with further questions such as age, place of residence, height, the number of shoes and the number of high-heel shoes in their possession, frequency of wearing high-heel shoes, the number of romantic and sexual partners. We asked them to evaluate their own level of attractiveness on a scale of 1 to 10, and to evaluate the level of attractiveness of the two men in the pictures used in the original study by Prokop and Švancárová (2020). We conduced focus groups in October and November 2020 during the Covid-19 pandemic. Due to government measures applicable at that time we had to conduct focus groups online using Microsoft Teams software.

Original pictures used as stimuli

With the permission of the authors of the original study Prokop and Švancárová (2020), we used their pictures of two men as stimuli, asking our participants what their first thoughts would be if the man in question had invited them on a date. We also asked about their general opinion of both men, and what they would wear to a date with him. The original authors found these pictures online, they were intended to represent an unattractive and an attractive man, respectively.

Data analysis

We used Thematic Analysis to analyze our qualitative data and followed the process suggested in their methodological paper titled "Using thematic analysis in psychology," and developed the approach in more depth in their 2021 book. The authors describe Thematic Analysis as a method for systematically identifying, organizing, and offering insight into patterns of meaning (themes) across a dataset. This method allows the researcher to see and make sense of collective or shared meanings and experiences. We chose this method because it allows for flexibility, permits the summarization of key properties in a dataset, and can generate unexpected knowledge. We recorded all of the focus group narratives and created transcripts from them. In the first stage we sought to identify the main themes relevant to our research questions. We then coded all the material and structured the codes hierarchically using Atlas.ti software: we assigned a code to every relevant unit of meaning that summarized its content, then we grouped coded into themes. The procedure observed all relevant ethical guidelines and identities of our subjects were not revealed.

Results

Our analysis identified four main themes in narratives regarding preparation for going on a date: (1) Young women see dates primarily as social events, rather than leading to sex; (2) High heels are not the primary means of beautification

for young women; (3) When dressing up for a date, young women consider practical aspects; and (4) Young women use beautification with caution.

Young women see dating as social events not necessarily leading to sex

When discussing dating it is vital to understand what participants understand under the term "date", where would this happen, how intimate such events could become. More than 52% of participants preferred dating in places such as restaurants, bars, cafes – often saying they felt safer in public spaces. When asked about a typical date, 46.2% would prefer going out (dinner, a wine bar, a beer pub; the legal age for drinking alcoholic beverages is 18 in Slovakia), and a further 28.8% envisaged a date as a walk or a trip. Associations were not primarily sexual; dates were more likely associated with enjoyable conversation and socialization. A date was generally considered to be a social event; here are two typical examples of describing dates:

> *V1: "We went to a restaurant, ordered dinner and just talked, it was all very cute."*

> *E1: "I associate my typical date with a dinner or a coffee and good conversation."*

When asked directly if they would consider having sex on a first date, 62% claimed they could not even imagine it. Participants refused sex on the first date mainly because they needed to get to know the partner and establish trust:

> *N1: "I cannot even imagine it; I need to get to know the person better and I would not feel comfortable if it happened."*

A further 36% explained that they could imagine having sex on the first date under certain circumstances, emphasizing that it would depend on how the date developed.

Overall, it seems that young women see first dates more as social events and opportunities to "screen out" potential partners. The idea of having sex on the first date seems to be remote to this age group. About a third admits it could happen, but only under the right circumstances.

High heels are not the primary means of beautification for young women

Although high heels may be associated with sexuality and seduction in popular culture or pornography, young women do not seem to share this association. We analyzed associations and most of our participants (59.1%) associated high-heel footwear with social events, 12.1% with work or school, 12.1% with dating, 9.1% with informal socialization such as going out for a coffee or dancing, and 7.6% would wear them to any suitable occasion. As many as 39 out of our 50 participants reported only wearing high heels on special occasions, and these were almost exclusively formal events such as weddings, prom nights, formal celebrations, family events, even religious functions. Here is a typical response regarding the events when they had last worn their high-heel shoes:

> K4: *"I choose high heels mostly for events such as a family celebration, an anniversary, a wedding, a holy communion, simply whenever I have to dress up."*

As an example, Participant M1 provided a photograph (see **Figure 1**) which suggests that her shoe collection contains two pairs of high heel shoes. The black high heel shoes on the right were only worn twice; she bought them for a Christmas party at her workplace. The other pair of high heel shoes (purple) were only worn at a prom night several years previously. All her remaining pairs are flat, four pairs are sneakers. This photograph represents a typical shoe collection of a young woman in our sample.

Young women in our sample generally preferred sneakers as their footwear of choice. Some specific brands such as Vans or Converse were associated with prestige or value.

We asked about beautification using this phrasing: *"Imagine a person you have fancied for some time asks you to go out on a Friday night. What would you wear? What would be going on in your head when choosing your outfit?"* Participants mentioned applying make-up, using jewelry (necklace, earrings), curling their hair – but only 8 answers included high heel shoes as spontaneous association.

To sum up, the majority of young women in our sample do not wear high heel shoes on a regular basis, only on special occasions. The only exceptions were those who wore

FIGURE 1
Photograph of a shoe collection, Participant M1.

such footwear professionally or semi-professionally – such as professional dancers or models. Our sample included two young women who tend to wear high heels regularly.

When dressing up for a date, young women consider practical aspects

When dressing up for a date young women tend to make themselves more beautiful, but the decision-making process includes very practical issues. When asked *"How do you decide what to wear to a romantic date?"* our participants named a lot of practical things they usually consider, and here are the most frequent ones:

(1) **Location of the date** (32 answers): Do they ride public transport to get there? Will there be cobblestone paved streets to walk through? Is there a chance of an outdoorsy walk or a hike?
(2) **Weather** (25 answers): Temperature, expected rain- or snowfall, etc.
(3) **Coordination with the partner** (15 answers): The style of the partner, and in case of high heels, also the height difference.

In summary, when choosing high heels or flats most of the young women consider location, weather, and coordination with partner.

Young women use beautification with caution

Our focus group participants do not always tend to beatify, they often do the very opposite – they try to avoid undesirable attention from men and dress down for situations where they have little control over interactions (e.g., when using public transport), or when they expect a company of people other than their peers and friends. Even in the context of dating they may try to project a casual image or divert attention from their bodies. When asked *"How would you dress for a date with Man 1, if he asked you out on a date?"*, 22.6% of the participants said they would not think too much about their clothes, and the most frequent answer was: jeans. With the more attractive Man 2, 19.3% would wear jeans, so there was very little difference (although 38.8% expressed they would put more thought and effort into their choice of outfit for a date with Man 2). Mean level of attractiveness of Man 1 was reported as $M = 2.88$ (SD = 2.23) while the attractiveness of Man 2 was $M = 7.5$ (SD = 3.73).

For example, our Participant K2 told us that on her first dates she tends not to show too much skin and tries to dress casually until she finds out more about the man and his intentions regarding a potential relationship. Another participant, L2 described jeans as "nicer clothes" when describing a hypothetical date with the "attractive" gentleman on the photograph:

> *"I would dress up nicer, differently than with the previous guy. (. . .) I would put on blue jeans and a nicer t-shirt or a blouse with ornaments, something more special. And sneakers in summer and boots with flat heels in winter."*

It seems that young women do use beautification consciously, but make sure to practice caution to avoid coming across as too eager or attracting undesirable attention to their bodies.

Discussion

As pointed out by Skipper and Nass (1966), dating fulfils a range of different functions. Respondents in our focus groups did not see dates as events leading to eventual reproduction, none of them expressed their desire to start a family anytime in the foreseeable future. They also did not see a date as an event that is necessarily linked with sex. The dating behavior of young women in our sample seemed more in line with observations by McDaniel (2005) – they searched for a combination of attractiveness and other positive personality traits. Although they seem to employ "behavioral tactics" (cf. Prokop and Švancárová, 2020) such as beautification, they use it with caution to avoid appearing too eager or projecting the wrong image.

Although high heels may be associated with sexuality in the pornographic imagery (Dietz and Evans, 1982), our respondents do not appear to use them as a common means of beautification. High heel shoes are much more associated with formal events and the need to comply with social norms. Moreover, only two out of 50 participants wear high heel shoes regularly, similarly to Prokop and Švancárová (2020) who reported that only two women in their sample of 292 young women claimed to wear high-heeled shoes eleven to twelve times per month and only one reported wearing shoes with high heels more than 25 times per month. Out of our 50 participants, 59.1% associate high-heel footwear with social events and reported wearing high heels on formal occasions such as weddings, prom nights, celebrations, etc. Although Wade and al. suggest high heels may trigger expectations of higher status, our participants would be more likely to wear sneakers of premium brands to demonstrate status in informal setting. Moreover, they would consider a range of other factors when selecting their outfit and footwear, primarily the location and associated practical issues, secondly the weather, and thirdly, coordination with their partner, which also seems to be a significant factor in determining the choice of outfit. This confirms the observation by Barnish et al. (2018) who pointed out that no study assessed the role of the respective heights of men and women in this

regard: many of our participants mentioned the height of the man as one of the important factors when deciding about the height of their heels.

Our focus group interviews suggest that young women do not seek a father of their future children when dating. Dating is mostly thought of as a social activity and does not necessarily lead to sex. Moreover, young women in our focus groups never brought up the topic of reproduction; having children is something that should be done after finishing college. In Slovakia this generally means three years of Bachelor training and 2 years of training for a Master?s degree. The participants in the original study as well as in our study were mostly still in the early years of their Bachelor studies.

There is a range of psychology studies that disprove the notion of rationality of partner choice. The summary of conceptualizations of love (Masaryk, 2012) suggests different neurobiological processes for sexual desire (mediated by gonadal estrogens and androgens) and for romantic love (mediated by reward systems such as endogenous opioids, catecholamines and neuropeptides such as oxytocin). There are, however, many other elements that may come into play. For example, one of the most cited authors in this regard, Robert Sternberg and Barnes (1988), later added a narrative component to his triangular theory of love. Couples whose narratives are complementary (not necessarily matching!) are the ones who tend to report satisfaction with their relationships. Another example would be Helen Fisher (2009) taxonomy of four groups (Explorers, Builders, Directors, Negotiators) tested on a sample of 7 million questionnaires. These studies suggest that partner choice is much more complicated than just psychical attractiveness.

When we introduced the two images to replicate the hypothetical scenario, the women asked the interviewers many additional questions to get a better understanding of the person. A mere photographic image appeared not to contain enough information for them to make a choice regarding the attractiveness of the man. Although they understood that one of the partners appeared to correspond to the cultural standard of attractiveness more than the other, they were still not sure whether he would make a suitable partner for a date. Many of our participants feared that the "attractive" man would be too intimidating, too promiscuous, unwilling to seek commitment, or simply not the kind of person with whom they would feel comfortable.

Our findings suggest that introducing a hypothetical scenario with two different young men does not necessarily lead young women to choose one as the preferred sexual partner. Although they may give such answer in a forced choice questionnaire, they usually hesitated to answer it in our focus groups and asked several questions to more fully understand the context of the date and the background of the man. As explained above, they seemed to understand the appeal of the "attractive" man, and understood this would be the culturally approved choice of the two, but they expressed many negative associations with the person in the picture. On the other hand, the "unattractive" man was viewed by some participants as cute, more successful, introverted, and less dominant. He was clearly not seen as someone who would be automatically disqualified from being a dating choice based on his appearance. They admitted they would wear nicer clothes to a date with Man 2, but we argue this would be more of an attempt to match his higher perceived social status than an attempt to beautify themselves in order to increase their mating chances. Although the authors had the images rated by fourteen women volunteers (who were older than the sample, at the mean age M = 25; SD = 1.74), our participants did not see the men as two extreme poles on the binary scale of attractive/unattractive. Nevertheless, they understood that one of them was more dominant and taller, and the other more submissive and shorter. Hence, if pressed, they would give answers corresponding to social convention. All this, on the other hand, does not entirely disprove the notion that high heel shoes are a strong sexual signal – but it seems they clearly do not serve this purpose for this age group of college undergraduates. Research involving older women with different experiences or from more deorived backgrounds might produce different results. A limitation of our study would obviously be the sensitive nature of the topic of dating and sexuality. Although we tried to create a very safe and trustful environment with two female researchers who were only marginally older than the participants, there is definitely a certain level of social desirability involved. We nevertheless believe that our study brings a more in-depth level of scrutiny, as well as more subtle and nuanced interpretation to the issues involved.

Conclusion

In conclusion, greater sophistication is needed beyond accepting simplistic assumptions about the way that women are driven to seek the most attractive partner to increase their reproductive chances. Humans clearly do not engage in sex solely to reproduce, and they do not go on dates only to have sex. Beautification or putting effort into choosing clothes may not necessarily relate to the drive for reproductive success and it may not be in any way linked to engaging in sexual intercourse. We argue, therefore, that using evolutionary theories as the basis for research on human behavior is not useful; and we advocate a shift to theories based on psychological, sociological and cultural factors.

Our qualitative take on the issue chosen by Prokop and Švancárová (2020) casts real doubt on the assertion that wearing high heels is a female mating strategy. Even in their own research, 38% of women reported never wearing high heels, 45% reported wearing them once per month, and the remaining 17% reported wearing them more than two times per month. This alone would question the notion of wearing high heels

being a regular part of dating behavior. Our data using photo solicitations suggests these numbers may be greatly exaggerated and young collegiate women in this age group wear high heel shoes far less frequently.

It is clear that young women do use beautification when preparing for a date. Specific forms however are likely going to be highly culture-specific and largely depend on the socio-economic status, cultural expectations, and identification with sub-cultures or fashion trends. Our sample and that collected by Prokop and Švancárová (2020) were from a Central European setting; it would be interesting to compare this with samples from other regions and look into possible relations with cultural values such as masculinity vs. femininity (cf. Bašnáková et al., 2016).

More and better research is needed to gain insight into decision-making by young people when it comes to dating; such research should, however, respect the diversity of social and cultural mores that influence the subtle choices being made. The enormous complexity of social behavior in humans makes it problematic to apply the rigid rules observed in animal mating, and concepts like parental investment may not be applicable to humans in the era when control over biological consequences is regarded as social norm while the imperative to reproduce does not justify breaking this norm. We need to be more attentive to social and psychological aspects when studying human behavior and avoid using science to reify stereotypes which do not accurately portray the reality of everyday life.

Data availability statement

The datasets presented in this article are not readily available because of ethical and privacy restrictions. Requests to access the datasets should be directed to the corresponding author.

Ethics statement

The studies involving human participants were reviewed and approved by Local Institutional Review Board, Faculty of Social and Economic Sciences, Comenius University Bratislava. The patients/participants provided their recorded informed consent to participate in this study. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

Author contributions

RM designed the study. NS and MB collected data. All authors provided substantial contribution to analyzing and interpreting data for the work, drafting the work, and provided approval for publication of the content.

Funding

This work was supported by VEGA Grant No. 1/0641/19.

Acknowledgments

We would like to thank Wendy Stainton Rogers for consultations regarding the language and style of the final manuscript.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

American College Health Association. (2018). *American College Health Association-National College Health Assessment II: Reference Group Executive Summary Fall 2018*. Silver Spring: American College Health Association.

Barnish, M., Morgan, H. M., and Barnish, J. (2018). The 2016 HIGh Heels: health effects And psychosexual BenefITS (HIGH HABITS) study: systematic review of reviews and additional primary studies. *BMC Public Health* 18:37. doi: 10.1186/s12889-017-4573-4

Bašnáková, J., Brezina, I., and Masaryk, R. (2016). Dimensions of culture: the case of Slovakia as an outlier in Hofstede's research. *Československá Psychologie* 60, 13–25.

Bjorklund, D., and Shackelford, T. (1999). Differences in Parental Investment Contribute to Important Differences Between Men and Women. *Curr. Direct. Psychol.* 8, 86–89. doi: 10.1111/1467-8721.00020

Buss, D. M. (1989). Sex Differences in Human Mate Preferences: evolutionary Hypotheses Tested in 37 Cultures. *Behav. Brain Sci.* 12, 1–49. doi: 10.1017/S0140525X00023992

Buss, D. M. (1994). *The Evolution of Desire: Strategies of Human Mating*. New York: Basic Books.

Dietz, P. E., and Evans, B. (1982). Pornographic imagery and prevalence of paraphilia. *Am. J. Psychol.* 139, 1493–1495. doi: 10.1176/ajp.139.11.1493

Eibl-Eibesfeldt, I. (1989). *Human Ethology*. Berlin: Aldine de Gruyter.

Fisher, H. (2009). *Why Him? Why Her?*. New York, NY: Oneworld.

Geary, D. C. (2000). Evolution and proximate expression of human paternal investment. *Psychol. Bull.* 126, 55–77. doi: 10.1037/0033-2909.126.1.55

Hewlett, B. S. (1992). "Husband-wife reciprocity and the father-infant relationship among Aka pygmies," in *Father-child relations: Cultural and biosocial contexts*, ed. B. S. Hewlett (New York: Aldine de Gruyter), 153–176.

Hills, R. (2015). *The sex myth: The gap getween our fantasies and reality*. New York, NY: Simon & Schuster.

Masaryk, R. (2012). Conceptualizing love: is it all we need? *Hum. Aff.* 22, 263–269. doi: 10.2478/s13374-012-0022-3

McDaniel, A. K. (2005). Young Women's Dating Behavior: why/Why Not Date a Nice Guy? *Sex Roles* 53, 347–359. doi: 10.1007/s11199-005-6758-z

Morris, P. H., White, J., Morrison, E., and Fisher, K. (2013). High heels as supernormal stimuli: how wearing high heels affects judgements of female attractiveness. *Evol. Hum. Behav.* 34, 176–181. doi: 10.1016/j.evolhumbehav.2012.11.006

Prokop, P., and Švancárová, J. (2020). Wearing high heels as female mating strategy. *Personal. Individ. Diff.* 152:109558. doi: 10.1016/j.paid.2019.109558

Puts, D. A. (2010). Beauty and the beast: mechanisms of sexual selection in humans. *Evol. Hum. Behav.* 31, 157–175. doi: 10.1016/j.evolhumbehav.2010.02.005

Skipper, J. K., and Nass, G. (1966). Dating Behavior: a Framework for Analysis and an Illustration. *J. Fam. Stud.* 28:412. doi: 10.2307/349537

Sternberg, R. J., and Barnes, M. L. (1988). *The Psychology of Love*. New Haven: Yale University Press.

Trivers, R. L. (1972). "Parental Investment and Sexual Selection," in *Sexual Selection and the Descent of Man*, ed. B. Campbell (Venice: Aldine), 1871–1971.

Wade, T. J., Burch, R. L., Fisher, M. L., and Casper, H. (2022). On a pedestal: high heels and the perceived attractiveness and evolutionary fitness of women. *Personal. Individ. Diff.* 188:111456. doi: 10.1016/j.paid.2021.111456

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 28 October 2022
DOI 10.3389/fpsyg.2022.1003242

Check for updates

OPEN ACCESS

EDITED BY
Oliver Niebuhr,
University of Southern Denmark, Denmark

REVIEWED BY
Fuguang Bao,
Zhejiang Gongshang University, China
Xiaozhou Zhou,
Université du Québec à Montréal, Canada
Francisco Javier Cabrerizo Lorite,
University of Granada, Spain

*CORRESPONDENCE
Xiangrui Chao
chaoxr@scu.edu.cn

SPECIALTY SECTION
This article was submitted to Human-Media Interaction, a section of the journal Frontiers in Psychology

RECEIVED 26 July 2022
ACCEPTED 28 September 2022
PUBLISHED 28 October 2022

CITATION
Chen J, Wang H and Chao X (2022) Cross-platform opinion dynamics in competitive travel advertising: A coupled networks' insight.
Front. Psychol. 13:1003242.
doi: 10.3389/fpsyg.2022.1003242

COPYRIGHT
© 2022 Chen, Wang and Chao. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Cross-platform opinion dynamics in competitive travel advertising: A coupled networks' insight

Jia Chen[1], Haomin Wang[2] and Xiangrui Chao[3]*

[1]School of Business Administration, Faculty of Business Administration, Southwestern University of Finance and Economics of China, Chengdu, China, [2]School of Management Sciences, Southwestern University of Finance and Economics of China, Chengdu, China, [3]Business School, Sichuan University, Chengdu, China

Social media platforms have become an important tool for travel advertisement. This study constructs the bounded confidence model to build an improved cross-platform competitive travel advertising information dissemination model based on open and closed social media platforms. Moreover, this study examines the evolution process of group opinions in cross-platform information dissemination with simulation experiments. Results reveal that based on strong relationships, the closed social media platform opinion leaders better guide in competitive travel advertising and can bring more potential consumers to follow. However, being an opinion leader on an open social media platform will not result in more consumer following.

KEYWORDS

coupled network propagation, competitive travel advertising, opinion leaders, opinion guidance, simulation model

Introduction

With the growing popularity of social networks, many tourists like to share their travel experiences (Lin et al., 2017; Liu et al., 2018). These social media platforms, such as TripAdvisor, Weibo, WeChat, and Ctrip, promote trust fusion among users through sharing, recommendation, communication, and other elements (Kah and Lee, 2014; Pötzschke and Braun, 2017; De Vries et al., 2017; Göbel and Munzert, 2018). This weakens the purpose of business information, makes social media publicity more convincing, and improves the company's marketing efficiency (Hajli and Sims, 2015). For example, online travel reviews on social platforms, including reviews of hotels and restaurants, have become an essential source of information for consumers making travel plans, thus enhancing hotel and restaurant advertising communication effects (Zhang et al., 2016). Compared with traditional information dissemination channels, the number of users of social platforms is growing, and it has become the preferred method for users to receive travel information. Indeed, public opinion has emerged as a critical norm in the dissemination of social information (Zhang et al., 2018; Kim et al., 2019; Shmargad and Sanchez, 2020; Ju et al., 2022).

As the primary means of disseminating public opinion, social media platforms are an important channel for Internet users to obtain much information. Therefore, as the second-largest online advertising platform, social media has attracted about 40% of advertisers to increase their advertising budgets, formulate personalized advertising content, share and spread, and generate various advertising effects (Nasir et al., 2021). However, due to the complexity of the Internet and social networks, product information dissemination is not always ideal and productive. Users cannot accurately judge the authenticity and accuracy of information; hence, product information is not widely disseminated, thereby influencing the advertising dissemination (Zhang et al., 2018). Prior research discovered that in the process of disseminating advertising information on social media platforms, opinion leaders are those who can influence other consumers' attitudes toward products by sharing their experiences with products or services and prompting them to make purchases, such as the big V travel blogger on Weibo (Eck et al., 2011). Moreover, prior study has found that opinion leaders typically have reliable knowledge in a specific field, a certain social and economic position, and the ability to attract others (Lazarsfeld et al., 1968). In addition, opinion leaders have high exposure in information dissemination and can influence other social users (Rogers, 2003). Therefore, opinion leaders have a significant impact on consumers' purchasing decisions on social media platforms. Research shows that 49% of users will rely on product recommendations from opinion leaders, and 40% will eventually purchase products recommended by opinion leaders (Karp, 2016). Therefore, using opinion leaders on social media platforms to promote products and guide consumers' opinions has become an important market strategy for tourism businesses.

Moreover, previous studies have examined the impact of opinion leaders on users' opinions on a single social platform (Eck et al., 2011; Dong et al., 2017; Zhao et al., 2018). However, the connection in the real world is becoming increasingly complex, and various information, such as hotel and restaurant reviews, are already flowing in multiple networks. These networks no longer exist in isolation, but they are interdependent and linked by structural and dynamic characteristics. These coupled systems can be found exist in multiple social platforms. For example, users use WeChat, Ctrip, and other apps to exchange information and Weibo, TripAdvisor, and other apps to share information and communicate with others.

Obviously, online travelers can be active in multiple social networks, and they can access and exchange information through multiple physical and social networks that intersect. Therefore, information is now distributed across multiple coupled networks rather than a single-platform network. The information dissemination of the coupled network has become more complex as the functions of social platforms have been upgraded and the range of travel users has expanded. As an important node of information dissemination, travel users directly determine the impact of information dissemination in the network. Travel users in a coupled network will receive information from multiple social network platforms, thus broadening the scope of information dissemination. If information dissemination in the coupled network is not managed, it may result in an accelerated attenuation of information propagation, thus reducing the effect of travel marketing information propagation (Zhang et al., 2018). Therefore, according to the dissemination law of coupled network information, establishing a model based on actual characteristics and constructing a coupled dissemination system are important ways to explore multi-platform travel information dissemination.

To investigate the influence of coupled network opinion leaders on consumers, this study builds a propagation model of competitive travel advertising information in coupled networks and analyzes the guiding process of opinion leaders in coupled networks to consumers' opinions based on the Hegselmann–Krause (HK) model of dissemination. Therefore, this study contributes in the following ways. First, we more realistically simulate the coupling propagation process of travel information by analyzing the dissemination and network characteristics of the two networks. Second, we analyze the dissemination mechanism of travel competitive advertising information in coupled networks and make recommendations for information dissemination on multi-platform networks in this manner.

The rest of this paper is structured as follows. Section 2 provides theoretical context and opinion dynamics. Section 3 develops an integrated public opinion dynamics model to examine the opinion evolution law of the coupled network's opinion leaders and followers. Section 4 shows the dissemination effect of competitive travel advertising in the coupling network under opinion leaders with the computer simulation. Finally, Section 5 presents the conclusion and discussions.

Literature review

In this section, we present definitions and properties of opinion leaders, and analysis the influence of opinion leaders. Then we analyze the characteristics of information dissemination, in addition, we focus on the opinion model in information dissemination, summarize the dissemination models in a single platform and multiple platforms, and analyze the necessity of research on multi-layer network cross-layer dissemination.

The influence of opinion leaders

Previous studies have shown that opinion leaders have an important influence on social platforms (Chen et al., 2016, 2021; Zhao et al., 2018). Opinion leaders, who are also social media influencers, usually utilize their ability to be "trusted person" in social media to influence brand awareness and the purchase decisions of large consumers (Cheng et al., 2019; Hudders et al., 2020). Furthermore, opinion leaders are social media micro-celebrities with a large following and significant influence on their audiences. This position on social media enables them to

communicate the brand's marketing message and influence consumer opinions (Delbaere et al., 2020). In the dissemination of electronic word-of-mouth (eWOM), opinion leaders often have definite, unwavering target opinions; their purpose is to influence other followers' opinions, and they are not affected by followers in the opinion update process (Zhao et al., 2018). In the opinion dynamics, opinion leaders, with more power, expertise, and positions, can affect other agents' opinions and achieve consensus or polarization of group decision-making (Dong et al., 2017; Zhang et al., 2020). Because opinion leaders' extensive exposure to mass media and close ties to change agents, which makes He/she becomes an influential social participant (Rogers, 2003). Therefore, the flow of public opinion and information is transferred from the mass media to the general public through the mediating role of opinion leaders (Lazarsfeld et al., 1944; Katz, 1957). From the perspective of opinion dynamics theory, according to the network structure, opinion leaders can play a key role in the network, most likely to influence the information flow of a large number of followers (Das et al., 2014). Simultaneously, several opinion dynamic models have been established and various experiments have been conducted to investigate the role of opinion leaders in the evolution of public opinion (Zhao and Kou, 2014; Chen et al., 2016, 2021). Opinion leaders will gradually shift public opinion to the desired target through micro-interaction during the opinion evolution process. Especially when they have similar opinions, gradually and intentionally changing others' opinions in the desired directions becomes easier (Afshar and Asadpour, 2010; Fan and Pedrycz, 2016).

Information dissemination

Travel information, such as online travel reviews, is a crucial source of information for tourists and facilitate their travel decisions (Duverger, 2013). The dissemination of travel information is essentially disseminating public opinion. Much research progress has been made on public opinion dissemination in the social networks. The spread of public opinion and infectious diseases are similar; thus, many scholars used the infectious disease model to study the public opinion dissemination. For instance, Wang et al. (2019) built a discrete communication model to discuss the spread of public opinion using the infectious disease susceptible, infected, and recovered model. They combined the hedging effects of negative and positive information. Their findings show that, in disseminating public opinion, netizens, the media, and the government will continuously optimize their strategies based on their own interests and information feedback. Opinion dynamics models, when applied to information dissemination, primarily examine how individuals interact and update their opinions in social networks. The opinion dynamics model is mainly used to describe specific aspects of the social behavior of a number of individuals and to simulate how the opinions of a group of groups evolve over time (Castro et al., 2018). There have been various approaches to analyze the process of changing these opinions, based on given various assumptions in the process. Using the continuous opinion and discrete actions model, Martins (2008) examined the discrete behavior of individual opinion interactions and deeply explored the impact of interaction rules on opinion evolution. Meanwhile, other researchers have examined the evolution of public opinion using a variety of public opinion dynamics models, for example, the Voter model, the DeGroot model, and the HK model, the details are shown in Table 1. As can be seen from Table 1, the opinion dynamics of a single social network has been extensively studied, both in terms of formation and evolution and opinion consensus reaching process, and these studies provide in-depth insights into the evolution of descriptions' opinions in an isolated network. In fact, public opinion dissemination is an overly complex dynamic process, and describing the dissemination process clearly is difficult. Especially with the development of information technology, information is no longer disseminated on a single platform but cross-disseminated in multiple platforms, making the description of public opinion dissemination increasingly complicated. However, there are few studies on the dissemination of cross-disseminated in multiple platforms. For example, the cross-layer propagation of single information online and offline is analyzed through the HK model (Ding et al., 2017; Dong et al., 2021; Ju et al., 2022). Because the cross-layer communication of online platforms is faster and more common. Therefore, based on the HK model, this paper constructs a cross-layer coupling propagation model of tourism advertising on open social media platforms and closed social media platforms, and further analyzes the information propagation in multi-layer networks.

Construction of a coupled two-layer online social network

The purpose of this paper is to analyze the dissemination of tourism advertising information in a multi-layer coupled network. Studies have shown that social network structure has an important impact on the dissemination of information (Zhang et al., 2018; He et al., 2021). Therefore, this section analyzes the structure of online social media platforms, and builds a two-layer social network.

Analysis of online social network structure

The dissemination of information is closely related to the structure of social networks. Complex network theories and methods are widely used in dissemination dynamics (Li et al., 2015; He et al., 2021). Existing research has shown that network topologies, such as WeChat and Facebook, have a critical impact on public opinion dissemination (Zhang et al., 2018; Mandal et al., 2020; Li and Wang, 2022).

TABLE 1 Opinion dynamics models at different social networks.

Application fields	Research object	Basic models	References
Single social network	Opinion formation and evolution	Bounded Confidence Model and Extensions	Zhao and Kou (2014), Zhao et al. (2018), Li et al. (2020), Hou et al. (2021), Li et al. (2021), Zhan et al. (2022)
		Voter model	Klamser et al. (2017), Jiao and Li (2021)
		Degroot model	Jia et al. (2017), Castro et al. (2018), Zhou et al. (2020), Wu et al. (2022)
	Opinion consensus reaching process	Bounded Confidence Model and Extensions	Pérez et al. (2018), Zhang et al. (2019), Zhang et al. (2020), Zhang et al. (2021)
		DeGroot model	Ding et al. (2019), Li and Wei (2019), Liang et al. (2022)
		Voter model	Gastner et al. (2018), Herrerías-Azcué and Galla (2019)
Multiple social networks	Opinion formation and evolution	Bounded Confidence Model and Extensions (online and offline)	Ding et al. (2017), Dong et al. (2021), and Ju et al. (2022)
		Voter model	Diakonova et al. (2016)
		SIR model	Zhang et al. (2018) and Jankowski and Chmiel (2022)
		Cross-network propagation model	Li and Wang (2022)

There are numerous social media platforms available today, and their structures vary greatly. For example, WeChat, Weibo, Facebook, Twitter, Ctrip, and other platforms have vastly different user connections and usage frequency. Some social media platforms, such as Weibo and Ctrip, only require one-way contact between users to communicate, even if they are unfamiliar with each other. These social platforms' network structure can be considered to have weak relationship strength. Meanwhile, in some social media platforms, such as WeChat and Facebook, users can only make contact through mutual authentication, which means they can only add friends and exchange information through authentication. Users on such social media platforms have stronger relationships and a higher level of trust. Previous research has classified existing social media platforms into two major groups (Li and Wang, 2022).

Open social medias

The connection between users is built based on one-way authentication. The users here usually have various friends, shallow social relationships, and weak friendship. In open social media, users can freely establish interactive relationships, and the number of users connected is large. Like Weibo, open social media users can forward other' Weibo content to their own Weibo through the forwarding function and like and reply to the content. Moreover, users usually have strong flexibility, which can be a one-way or two-way relationship. A weak relationship network of "radiation" is formed through the user's attention, which triggers the "secondary radiation propagation" of information. The structure of open social media is the directed, scale-free networks (Barabasi and Albert, 1999). Therefore, this paper chooses BA-directed scale-free networks to simulate these social media platforms.

Closed social medias

The establishment of the connection relationship between users is achieved through mutual authentication. Usually, the number of users' friends is small, and the mutual trust is higher. The addition of user friends on a closed social media platform is mostly recommended by other users or searched for by the system. Thus, the growth of this network structure is random. Users add friends through mobile phone numbers, QQ friends, and so on, similar to WeChat, thus forming a peer-to-peer communication mode. This is a social model based on offline acquaintances who have a strong bond with one another. The majority of group communication occurs through the formation of WeChat groups, and the information in the user's circle of friends can be seen by other friends, resulting in information group communication. These communication methods rely on a "circle" network with strong relationships. Additionally, information dissemination has privacy, and individuals have a higher degree of trust. The structure of these social media is a typical undirected BA scale-free network (Traud et al., 2012). Additionally, this paper chooses an undirected BA scale-free network to simulate these social media platforms.

Construction of two-layer coupling online social networks

Social diversity has become a defining feature with the new media development. This means that public opinion spreads not only on one social media platform but also across networks in multiple platforms. For example, users may capture Ctrip travel information and forward it to WeChat; similarly, WeChat travel experiences may be forwarded to Twitter. To explore cross-platform information dissemination, we define two online social networks, namely, network A and network B. Network A is a closed social media platform, whereas network B is an open social media platform. In Figure 1, the edges between nodes in each layer represent social relationships in the platform. The connecting edge between networks A and B indicates that a user can have

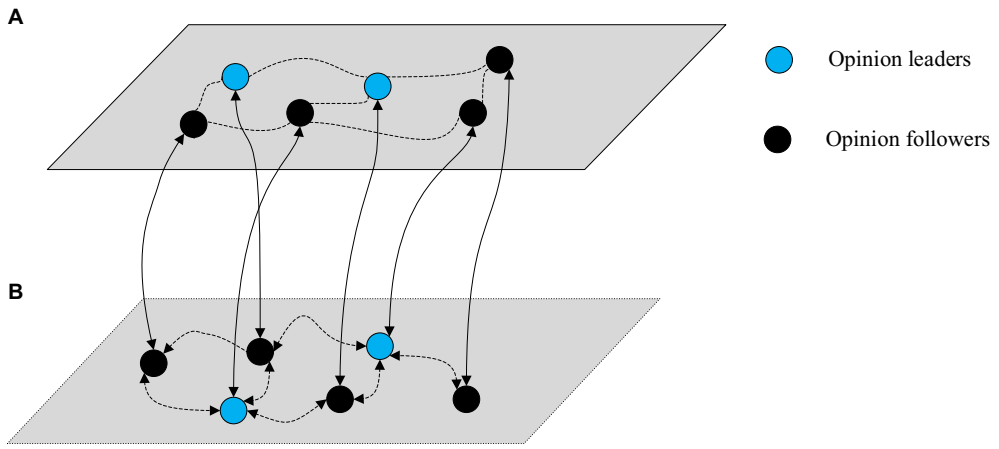


FIGURE 1
(A,B) Two-layer coupled network structure.

accounts in multiple social networks. Furthermore, in this coupled network, the correspondence between networks A and B is one-to-one; others are ignored (Li and Wang, 2022). Nodes in networks A and B are connected at random. In addition, the opinion leader in network A may be a follower in network B. Similarly, an opinion leader in network B could be an opinion follower in network A.

In this coupled network, the following are some assumptions:

1. Each user has one and only one account in networks A and B. This means that individuals on open social media platforms and closed social media platforms can receive any information from both platforms at the same time.
2. The user's addition or deletion is not considered; that is, the network is static.
3. The states of the same node in two different network layers are allowed to be different.
4. A group of opinion leaders exists in networks A and B, and their status is affected only by the opinions of the target travel advertisement;
5. Once the opinion leader group publishes an opinion, networks A and B users can receive the message immediately.

Design of competitive advertising propagation model in coupled network

In social media platforms, potential consumers will always trust individuals with similar opinions (Liu et al., 2015). Therefore, the HK model can better describe the dynamic evolution process of consumer opinions. The original HK model is defined as:

Let $X(t)=\{x_1(t),x_2(t),\ldots x_N(t)\}$ be the set of opinions of individual i at time t. For the case $|x_i(t)-x_j(t)|\leq\varepsilon$, the opinion update rule of individual i at time $t+1$ is as follows (Heselmann and Krause, 2002):

$$x_i(t+1)=\frac{\sum_{j:|x_i(t)-x_j(t)|\leq\varepsilon}\sigma_{ij}x_j(t)}{\sum_{j:|x_i(t)-x_j(t)|\leq\varepsilon}\sigma_{ij}} \quad (1)$$

where, ε is the bounded trust level of the individual, and σ_{ij} is the weight that individual i assigns to the individual j at time.

Under the framework of bounded trust theory, we construct two competitive opinion groups in a two-layer coupled network to analyze the evolution process of individuals. Without loss of generality, this paper assumes an opinion leader group exist in both networks A and B. Each opinion leader group represents opposing advertising opinions for competitive products. Therefore, the target advertising opinion of the opinion leader group in network A is 1, whereas the target travel advertising opinion of the opinion leader group in network B is −1.

The dissemination of advertising opinions depends on the network's topology (Boccaletti et al., 2006). Assuming that $[a_{ij}]_{N\times N}$ is the adjacency matrix of network A, we determine that $a_{ij}=1$ means a connection exists between individuals i and j in network A; otherwise, $a_{ij}=0$ denotes no connection between individuals i and j. Meanwhile, $[b_{ij}]_{N\times N}$ is the adjacency matrix of network B: if $b_{ij}=1$, a connection exists between individuals i and j in network B; otherwise, $b_{ij}=0$ denotes no connection. If individuals i and j are connected, individual i can receive an opinion from individual.

According to the HK model, if $|x_i(t)-x_j(t)|\leq\varepsilon$, the rule of opinion leader groups in network A is defined as:

$$x_i^{L_1}(t+1)=(1-w_1)\left[\frac{p_i}{N_i^{L_1}(t)}\sum_{j=N_1+1}^{N_1+N_2}\delta_{ij}^1(t)a_{ij}x_j(t)+(1-p_i)x_i(t)\right]+w_1d_1 \quad (2)$$

where, $i=N_1+1,\ldots,N_1+N_2$, $\delta_{ij}^1(t)=\begin{cases}0, otherwise\\ 1, \|x_j(t)-x_i(t)\|\leq\varepsilon_1\end{cases}$,

ε_1 is the bounded confidence level of the individuals in the network A. $\frac{1}{N_i^{L_1}(t)}=\sum_{j=N_1+1}^{N_1+N_2}\delta_{ij}^1(t)a_{ij}$ represents the number of neighbors

of opinion leaders in network A. As aforementioned, a_{ij} represents the connection of network A, which is 0 or 1. w_1 is the influence weight of the target advertisement in the network A. p_i represents the level to which the individual is affected by the opinion from the network A. Therefore, $1 - p_i$ is the individual's self-confidence degree when receiving the opinion, and d_1 is the target travel advertisement opinion value in the network A. This shows that the opinions of opinion leaders are mainly influenced by targeted advertisements and other opinion leaders in the same layer group.

As aforementioned, the rule of opinion leader groups in network B is defined as follows:

$$x_i^{L_2}(t+1) = (1-w_2)\left[\frac{q_i}{N_i^{L_2}(t)}\sum_{j=N_1+N_2+1}^{N}\delta_{ij}^2(t)b_{ij}x_j(t) + (1-q_i)x_i(t)\right] + w_2 d_2 \quad (3)$$

where, $i = N_1 + N_2 + 1,\ldots,N$, $\delta_{ij}^2(t) = \begin{cases} 0, otherwise \\ 1, \| x_j(t) - x_i(t) \| \le \varepsilon_2 \end{cases}$,

ε_2 is the bounded confidence level of the individuals in the network B. $\frac{1}{N_i^{L_2}(t)} = \sum_{j=N_1+N_2+1}^{N}\delta_{ij}^2(t)b_{ij}$ represents the number of neighbors of opinion leaders in network B. b_{ij} represents the connection of network B, which is 0 or 1. w_2 is the influence weight of the target advertisement in network B. q_i represents the level to which the individual is affected by the opinion from network B. Therefore, $1 - q_i$ is the individual's self-confidence degree when receiving the opinion, and d_2 is the target travel advertisement opinion value in the network B.

The opinion update model for opinion followers in networks A and B is:

$$\begin{aligned} x_i^F(t+1) = p_i[&\alpha\frac{1}{N_i^{L_1}(t)}\sum_{j=N_1+1}^{N_1+N_2}\delta_{ij}^1(t)a_{ij}x_j(t) \\ &+(1-\alpha)\frac{1}{N_i^{F_1}(t)}\sum_{j\in F_1}\delta_{ij}^1(t)a_{ij}\left(\begin{array}{l} q_i\frac{1}{N_j'^{F_2}(t)}\sum_{k\in F_2}\delta_{jk}'(t)b_{jk}x_k(t) \\ +(1-q_i)x_j(t)\end{array}\right)] \\ &+q_i[\alpha\frac{1}{N_i^{L_2}(t)}\sum_{j=N_1+N_2+1}^{N}\delta_{ij}^2(t)b_{ij}x_j(t) \\ &+(1-\alpha)\frac{1}{N_i^{F_2}(t)}\sum_{j\in F_2}\delta_{ij}^2(t)b_{ij}\left(\begin{array}{l} p_i\frac{1}{N_j''^{F_1}(t)}\sum_{k\in F_1}\delta_{jk}''(t)a_{jk}x_k(t) \\ +(1-p_i)x_j(t)\end{array}\right)] \\ &+(1-p_i-q_i)x_i^F(t) \end{aligned} \quad (4)$$

where, $\delta_{jk}''(t) = \begin{cases} 0, otherwise \\ 1, \| x_j(t) - x_k(t) \| \le \theta_1 \end{cases}$,

$\delta_{jk}'(t) = \begin{cases} 0, otherwise \\ 1, \| x_j(t) - x_k(t) \| \le \theta_2 \end{cases}$, θ_1 and θ_2 represents the inter-layer propagation thresholds of networks A and B. That is, the threshold for network A users to spread their opinion to network B after accepting their opinion is θ_1, whereas the threshold for network B users to spread their opinion to network A after accepting their opinion is θ_2. Target advertising opinions spread cross-network in coupled networks A and B when the values of θ_1 and θ_2 are less than or equal to the bounded confidence level, respectively. α is the degree of followers affected by the group of opinion leaders in networks A and B. F_1 and F_2 represent the set of followers of the opinion leaders' opinion dissemination across layers in networks A and B, where $F_1 = \{1,\ldots,N_1\} \cup \{N_1 + N_2 + 1,\ldots,N\}$, $F_2 = \{1,\ldots,N_1 + N_2\}$, and

$N_i^{F_1}(t) = \sum_{j\in F_1}\delta_{ij}^1(t)a_{ij}$, $N_i^{F_2}(t) = \sum_{j\in F_2}\delta_{ij}^2(t)b_{ij}$,

$N_j'^{F_2}(t) = \sum_{k\in F_2}\delta_{jk}'(t)b_{jk}$, and $N_j''^{F_1}(t) = \sum_{k\in F_1}\delta_{jk}''(t)a_{jk}$.

Simulation analysis of competitive advertising in coupling networks

The opinion dynamics model usually describes the evolution of group opinions with a simulation (Castro et al., 2018), therefore, this study uses a computer simulation method to analyze the group opinions dynamic evolution process of opinion leaders who promote competitive travel advertisements in coupled network. Some initial assumptions are applied in the experiments:

1. The coupled network has 1,000 nodes: networks A and B have 10 opinion leaders, respectively, and the remaining nodes are followers.
2. The initial opinions of opinion leaders and followers all obey the uniform distribution on [−1,1].
3. The confidence levels of individuals in networks A and B are $\varepsilon_1 = \varepsilon_2 = 0.5$; the level of individuals affected by the opinions of others in networks A and B is $q_i = q_i = 0.5$; followers in networks A and B are influenced by opinion leaders $\alpha = 0.5$.
4. The threshold for network A (B) users to spread the opinion to network B (A) after accepting opinion is $\theta_1 = \theta_2 = 0.3$.
5. The opinion values of target travel advertisements in networks A and B are $d_1 = 1$, $d_2 = -1$; the weight of target travel advertisements is $w_1 = w_2 = 0.5$.

As shown in Figure 2, the red, green, and blue lines represent opinion leader group in network A, opinion leader group in network B, and opinion followers, respectively. When the same travel advertising weight influences the two network platforms, number of opinion leaders, and confidence level, the opinion leaders in the two networks quickly converge to the target travel advertising opinions, and the followers' opinions quickly converge to the middle opinion value of 0. This shows that under the influence of factors, such as the same travel advertising intensity, potential followers will not appear to be biased toward a certain network of travel advertising opinions.

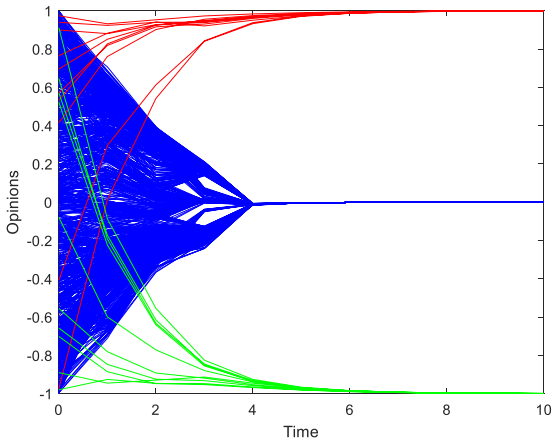


FIGURE 2
The opinion evolution with initial conditions.

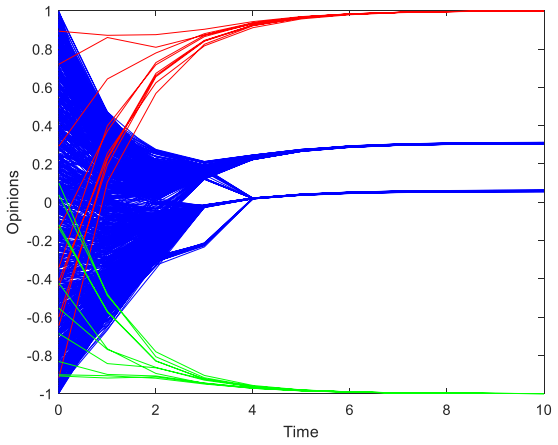


FIGURE 3
Opinion evolution with changes in confidence level.

However, in a closed social network, individuals usually have a strong relationship with each other than in open social network, so they usually have a higher level of trust with each other (Li and Wang, 2022) and a higher level of confidence in others. Therefore, assume that $\varepsilon_1 = 0.7$, $\varepsilon_2 = 0.5$ and other parameters are as above.

Figure 3 shows that, in a coupled network, followers' opinions eventually converge in the intervals [0.2, 0.4], and [0, 0.2], indicating that followers' opinions generally tend to target travel advertising of closed social media while completely ignoring open social media advertising. This demonstrates that, when all other conditions remain constant, travel advertisements in closed social media are more likely to be accepted by potential consumers than open social media, thus bringing more potential consumers to follow in both social media platforms.

Closed social media travel advertisements can bring more potential consumers to follow. How then can open social media companies take measures to further enhance potential consumers' recognition of their advertisements? The first measure is to increase the open social media travel advertising weight.

In Figure 4, the confidence level is $\varepsilon_1 = 0.7$, $\varepsilon_2 = 0.5$, the weight of advertisement is $w_1 = 0.3$, $w_2 = 0.9$, and other parameters are as aforementioned. The results reveal that with the increase in travel advertisement weight, the opinions of opinion leaders in network B quickly converge to the target travel advertisement opinion value of -1. However, followers' opinions eventually converge to an interval greater than 0. This demonstrates that in cross-platform communication, opening social media by increasing the weight of travel advertising will not result in followers recognizing the target travel advertisements. They continue to rely on closed social network travel advertising information. Therefore, increasing the weight of travel advertising is ineffective.

To analyze the impact of the number of opinion leaders on potential users in the coupled network, this study increases the number of opinion leaders on open social platforms. The results are shown in Figure 5. The confidence level is $\varepsilon_1 = 0.7$, $\varepsilon_2 = 0.5$, and the weight of the advertisement is $w_1 = 0.5$, $w_2 = 0.5$. The number of opinion leaders is $N_1 = 10$, $N_2 = 40$, $N_2 = 80$, $N_2 = 120$, respectively.

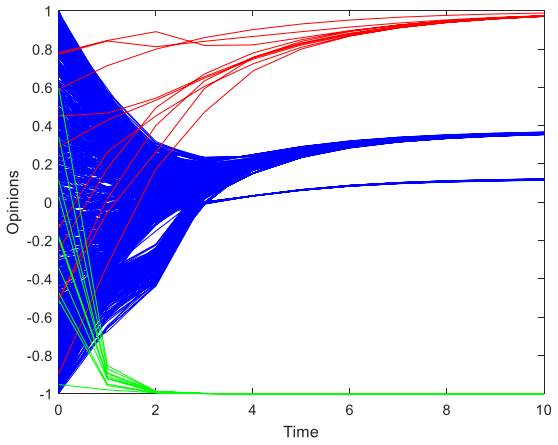


FIGURE 4
The opinion evolution with changes of travel advertisement weight.

Results reveal that as the number of opinion leaders on the open platform grows, followers' opinions gradually converge to the middle opinion value of 0, and they no longer only follow to the target travel advertisement of closed social platforms. However, as the number of opinion leaders grows, the opinion value of followers returns to the interval above 0.

This result demonstrates that a moderate increase in the number of opinion leaders on open platforms can appropriately guide the evolution of consumers' opinions on travel advertising in the coupled network. However, it cannot finally make consumers recognize the travel advertising of open social platforms.

Conclusion and discussion

This paper investigates the process by which potential consumers' opinions evolve in a coupled network under the influence of competitive travel advertising promoted by opinion

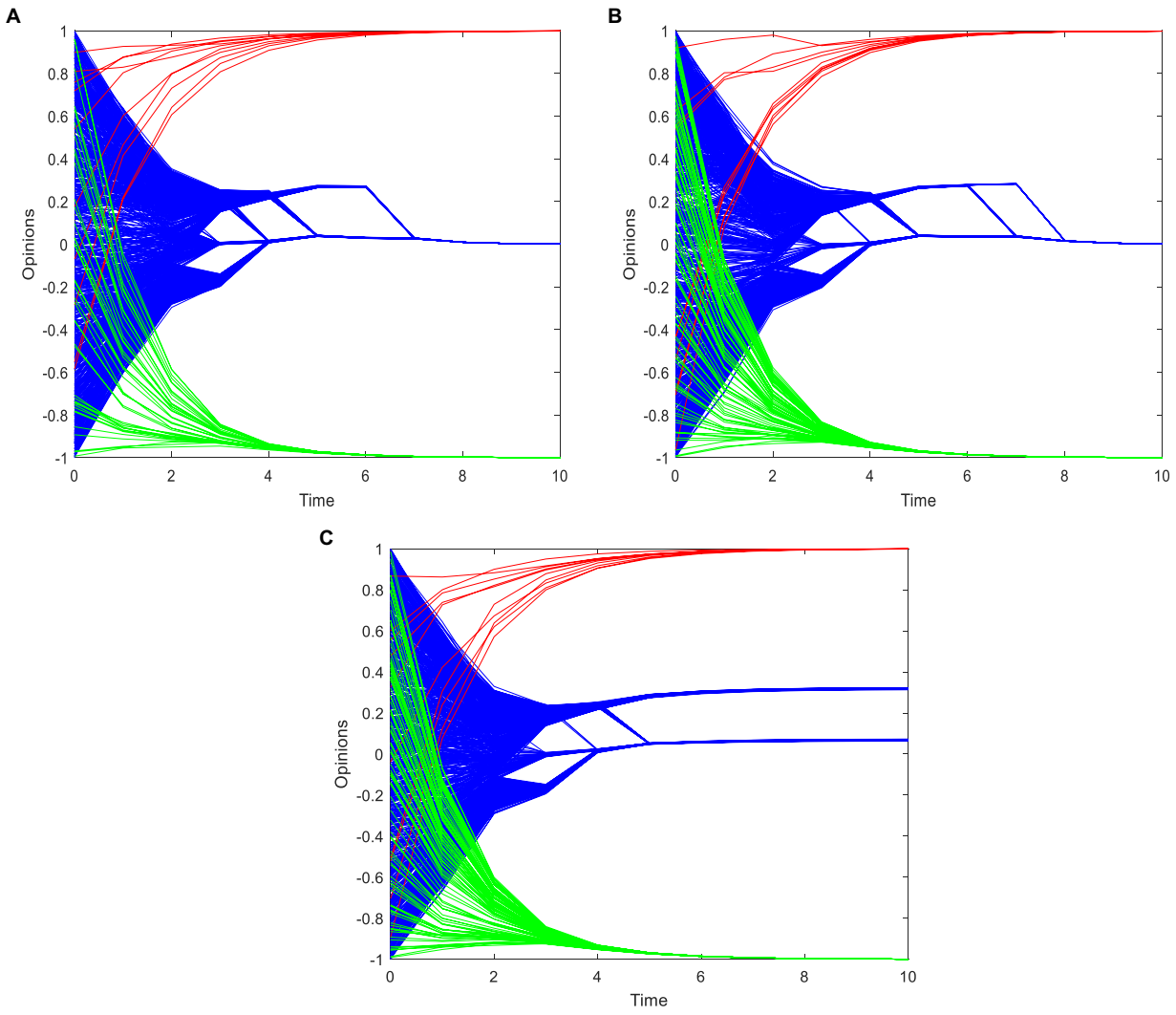


FIGURE 5
The opinion evolution with opinion leader. **(A)** N_1=10, N_2=40. **(B)** N_1=10, N_2=80. **(C)** N_1=10, N_2=120.

leaders. The evolution of consumers' opinions in real closed social media and open social media is simulated in a computer by building a bounded confidence opinion dynamics model of individuals in a cross-platform coupled network. This study provides a scientific strategy for travel advertising or WOM promotion on multiple social media platforms.

The results can be summarized as follows:

1. For multiple social media platforms, closed social media have better cross-platform guidance effects in the process of cross-platform dissemination of competitive travel advertisements, whereas open social media are less effective than closed social media. The findings of this study differ from the findings of the single-platform opinion evolution study (Zhao et al., 2018). In a single platform, the final followers' opinion is symmetrically distributed with opinion interval, and no apparent bias exists toward any one opinion leader subgroup. The results show that increasing the confidence level did not significantly improve the opinion leaders' influence. However, this study found that the final opinion of followers clearly favors opinion leaders in closed social media platforms in a coupled network. This demonstrates that in a coupled network, the followers' choice of leaders is influenced by their confidence level.
2. Because competitive travel advertisements are spread across the coupled platforms, increasing the weight of travel advertising will not improve the cross-platform guidance effect of advertising. This finding differs from that of previous studies (Previte, 1999; Coker, 2017). Prior research has found that competing advertisement influence should be in an effective range; otherwise, the advertisement will suffer a negative effect (Previte, 1999; Coker, 2017). However, in the coupled network,

increasing the influence weight of travel advertising will not result in more cross-platform consumers following. The possible reason is that in cross-platform information dissemination, consumers are more cautious about advertisements and more prone to question the advertisement information (Priester and Petty, 1995; Yang and Hsu, 2017). Although the influence weight of travel advertisements has increased, influencing consumers' decision-making is not enough.

3. Opinion leaders of open social media play a lesser role in cross-platform information dissemination than closed social network platforms. Previous studies have shown that opinion leaders on open social platforms have a critical influence on the scale of information dissemination on a single platform (Luqiua et al., 2019; Wang et al., 2020). However, this paper found that, in the coupled network, although the open platform opinion leaders will affect the potential consumers' opinions, it ultimately failed to make potential consumers to follow. Therefore, in the coupled network, the effect of open social media opinion leaders is less evident than signal-platform.

This study also has some practical implications. Companies should focus on increasing consumer trust in open social media platforms. The closed social media platform is based on offline social relationships, and friends' connections and so on generally have a higher level of trust. Therefore, when travel information from multiple platforms affects consumers simultaneously, potential consumers are more likely to choose travel advertising information on a closed social media platform with higher trust over open social media with lower trust. Therefore, open social platforms should focus on improving social platform trust relationships, thereby increasing the cross-platform dissemination effect of open platforms. Simultaneously, for a better cross-platform publicity effect on the opening platform, the number of opinion leaders in travel advertising can be appropriately increased. However, how to determine the number of opinion leaders deserves further exploration.

At the same time, this paper also has some limitations. For example, this paper focuses on the difference in propagation properties between the opinion leader and the follower. In the future research, the properties of the opinion leader and the follower can be further considered, such as hobbies, social status, etc (Anagnostopoulos et al., 2022), which can be closer to the information propagation process in the real environment. In addition, this paper only considers the cross-platform dissemination process of competing advertising information in two online platforms. In the future, we can consider building a cross-platform competitive dissemination model of advertising in online-offline multi-layer networks (Ju et al., 2022), so as to better describe the actual dissemination process of competing advertising information.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Author contributions

JC: conceptualization, formal analysis, and writing—original draft. HW: writing—review and editing. XC: conceptualization, methodology, and writing—review and editing. All authors contributed to the article and approved the submitted version.

Funding

This work was supported in part by grants from the National Social Science Foundation of China(#22BGL236), National Natural Science Foundation of China (#72274132), and the Fundamental Research Funds for the Central Universities (JBK2201026).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Afshar, M., and Asadpour, M. (2010). Opinion formation by informed agents, journal of artificial societies and social. *Simulation* 13:5. doi: 10.18564/jasss.1665

Anagnostopoulos, A., Becchetti, L., Cruciani, E., Pasquale, F., and Rizzo, S. (2022). Biased opinion dynamics: when the devil is in the details. *Inf. Sci.* 593, 49–63. doi: 10.1016/j.ins.2022.01.072

Barabasi, A., and Albert, R. (1999). Emergence of scaling in random networks. *Science* 286, 509–512. doi: 10.1126/science.286.5439.509

Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., and Hwang, D. U. (2006). Complex networks: structure and dynamics. *Phys. Rep.* 424, 175–308. doi: 10.1016/j.physrep.2005.10.009

Castro, J., Lu, J., Zhang, G., Dong, Y., and Martínez, L. (2018). Opinion dynamics-based group recommender systems. *IEEE Trans. Syst. Man Cyber. Syst.* 48, 2394–2406. doi: 10.1109/TSMC.2017.2695158

Chen, S., Glass, D. H., and McCartney, M. (2016). Characteristics of successful opinion leaders in a bounded confidence model. *Physica A* 449, 426–436. doi: 10.1016/j.physa.2015.12.107

Chen, J., Kou, G., Wang, H., and Zhao, Y. (2021). Influence identification of opinion leaders in social networks: an agent-based simulation on competing advertisements. *Inform. Fus.* 76, 227–242. doi: 10.1016/j.inffus.2021.06.004

Cheng, X., Fu, S., Sun, J., Bilgihan, A., and Okumus, F. (2019). An investigation on online reviews in sharing economy driven hospitality platforms: a viewpoint of trust. *Tour. Manag.* 71, 366–377. doi: 10.1016/j.tourman.2018.10.020

Cheng, C., and Yu, C. (2019). Opinion dynamics with bounded confidence and group pressure. *Physica A* 532:121900. doi: 10.1016/j.physa.2019.121900

Coker, B. (2017). The ineffectiveness of counterclaim advertising for increasing consumer sentiment. *J. Consum. Behav.* 16, 34–41. doi: 10.1002/cb.1600

Das, A., Gollapudi, S., and Munagala, K. (2014). Modeling opinion dynamics in social networks, in Proceedings of the 7th ACM international conference on web search and data mining, 403–412). doi: 10.1145/2556195.2559896.

De Vries, L., Gensler, S., and Leeflang, P. (2017). Effects of traditional advertising and social messages on brand-building metrics and customer acquisition. *J. Mark.* 81, 1–15. doi: 10.1145/2556195.2559896

Delbaere, M., Michael, B., and Phillips, B. (2020). Social media influencers: a route to brand engagement for their followers. *Psychol. Mark.* 38, 101–112. doi: 10.1002/mar.21419

Diakonova, M., Nicosia, V., Latora, V., and Miguel, M. S. (2016). Irreducibility of multilayer network dynamics: the case of the voter model. *New J. Phys.* 18:023010. doi: 10.1509/jm.15.0178

Ding, Z., Chen, X., Dong, Y., and Herrera, F. (2019). Consensus reaching in social network DeGroot Model: The roles of the Self-confidence and node degree. *Inform. Sci.* 486, 62–72. doi: 10.1016/j.ins.2019.02.028

Ding, Z., Dong, Y., Liang, H., and Chiclana, F. (2017). Asynchronous opinion dynamics with online and offline interactions in bounded confidence model. *J. Artif. Soc. Soc. Simul.* 20:6. doi: 10.18564/jasss.3375

Dong, Y., Ding, Z., Chiclana, F., and Herrera-Viedma, E. (2021). Dynamics of public opinions in an online and offline social network. *IEEE Trans. Big Data* 7, 610–618. doi: 10.1109/TBDATA.2017.2676810

Dong, Y. C., Ding, Z. G., Martínez, L., and Herrera, F. (2017). Managing consensus based on leadership in opinion dynamics. *Inf. Sci.* 397, 187–205. doi: 10.1016/j.ins.2017.02.052

Duverger, P. (2013). Curvilinear effects of user-generated content on hotels' market share: a dynamic panel-data analysis. *J. Travel Res.* 52, 465–478. doi: 10.1177/0047287513478498

Eck, P., Jager, W., and Leeflang, P. (2011). Opinion Leaders' role in innovation diffusion: a simulation study. *J. Prod. Innov. Manag.* 28, 187–203. doi: 10.1111/j.1540-5885.2011.00791.x

Fan, K., and Pedrycz, W. (2016). Opinion evolution influenced by informed agents. *Physica A* 462, 431–441. doi: 10.1016/j.physa.2016.06.110

Gastner, M. T., Oborny, B., and Gulyás, M. (2018). Consensus time in a voter model with concealed and publicly expressed opinions. *J. Stat. Mech. Theory Exp.* 6:063401. doi: 10.1088/1742-5468/aac14a

Göbel, S., and Munzert, S. (2018). Political advertising on the Wikipedia marketplace of information. *Soc. Sci. Comput. Rev.* 36, 157–175. doi: 10.1177/0894439317703579

Hajli, N., and Sims, J. (2015). Social commerce: the transfer of power from sellers to buyers. *Technol. Forecast. Soc. Chang.* 94, 350–358. doi: 10.1016/j.techfore.2015.01.012

He, Q., Sun, L., Wang, X., Wang, Z., and Ma, L. (2021). Positive opinion maximization in signed social networks. *Inf. Sci.* 558, 34–49. doi: 10.1016/j.ins.2020.12.091

Herrerías-Azcué, F., and Galla, T. (2019). Consensus and diversity in multistate noisy voter models. *Phys. Rev. E* 100:022304. doi: 10.1103/PhysRevE.100.022304

Heselmann, R., and Krause, U. (2002). Opinion dynamics and bounded confidence models, analysis and simulation. *J. Artif. Soc. Soc. Simul.* 5, 1–33.

Hou, J., Li, W., and Jiang, M. (2021). Opinion dynamics in modified expressed and private model with bounded confidence. *Physica A* 574:125968. doi: 10.1016/j.physa.2021.125968

Hudders, L., De Jans, S., and De Veirman, M. (2020). The commercialization of social media stars: a literature review and conceptual framework on the strategic use of social media influencers. *Int. J. Advert.* 39, 1–49. doi: 10.1080/02650487.2020.1836925

Jankowski, R., and Chmiel, A. (2022). Role of time scales in the coupled epidemic-opinion dynamics on multiplex networks. *Entropy* 24:105. doi: 10.3390/e24010105

Jia, P., Friedkin, N. E., and Bullo, F. (2017). Opinion dynamics and social power evolution over reducible influence networks. *SIAM J. Control. Optim.* 55, 1280–1301. doi: 10.1137/16M1065677

Jiao, Y., and Li, Y. (2021). An active opinion dynamics model: the gap between the voting result and group opinion. *Inform. Fus.* 65, 128–146. doi: 10.1016/j.inffus.2020.08.009

Ju, C., Wang, C., Jiang, Y., Bao, F., Zhou, H., and Xu, C. (2022). Exploring a multi-layer coupled network propagation model based on information diffusion and bounded trust. *Int. J. Public Health* 67:1604887. doi: 10.3389/ijph.2022.1604887

Kah, J. A., and Lee, S. H. (2014). Beyond adoption of travel technology: its application to unplanned travel behaviors. *J. Travel Tour. Mark.* 31, 667–680. doi: 10.1080/10548408.2014.888968

Karp, K. (2016) New research: the value of influencers on twitter, https://blog.twitter.com/2016/new-research-the-value-of-influencers-on-twitter.

Katz, E. (1957). The two-step flow of communication: an up-to-date report on an hypothesis. *Public Opin. Q.* 21, 61–78. doi: 10.1086/266687

Kim, J. J., Lee, Y., and Han, H. (2019). Exploring competitive hotel selection attributes among guests: an importance-performance analysis. *J. Travel Tour. Mark.* 36, 998–1011. doi: 10.1080/10548408.2019.1683484

Klamser, P., Wiedermann, M., Donges, J., and Donner, R. (2017). Zealotry effects on opinion dynamics in the adaptive voter model. *Phys. Rev. E* 96:052315. doi: 10.1103/PhysRevE.96.052315

Lazarsfeld, P. F., Berelson, B., and Gaudet, H (1968). *The people's choice*. How the Voter Makes Up His Mind in a Presidential Campaign, New York Chichester, West Sussex: Columbia University Press.

Li, Q., Du, Y., Li, Z., Hu, J., Hu, R., Lv, B., et al. (2021). HK–SEIR model of public opinion evolution based on communication factors. *Eng. Appl. Artif. Intell.* 100:104192. doi: 10.1016/j.engappai.2021.104192

Li, K., Liang, H., Kou, G., and Dong, Y. (2020). Opinion dynamics model based on the cognitive dissonance: an agent-based simulation. *Inform. Fus.* 56, 1–14. doi: 10.1016/j.inffus.2019.09.006

Li, D., Ma, J., Tian, Z., and Zhu, H. (2015). An evolutionary game for the diffusion of rumor in complex networks. *Physica A* 433, 51–58. doi: 10.1016/j.physa.2015.03.080

Li, Y., and Wang, J. (2022). Cross-network propagation model of public opinion information and its control in coupled double-layer online social networks. *Aslib J. Inf. Manag.* 74, 354–376. doi: 10.1108/AJIM-04-2021-0126

Li, S., and Wei, C. (2019). Modeling the social influence in consensus reaching process with interval fuzzy preference relations. *Int. J. Fuzzy Syst.* 21, 1755–1770. doi: 10.1007/s40815-019-00671-5

Liang, X., Guo, J., and Liu, P. (2022). A large-scale group decision-making model with no consensus threshold based on social network analysis. *Inf. Sci.* 612, 361–383. doi: 10.1016/j.ins.2022.08.075

Lin, P. M. C., Qiu Zhang, H., Gu, Q., and Peng, K. L. (2017). To go or not to go: travel constraints and attractiveness of travel affecting outbound Chinese tourists to Japan. *J. Travel Tour. Mark.* 34, 1184–1197. doi: 10.1080/10548408.2017.1327392

Liu, S., Jiang, C., Lin, Z., Ding, Y., Duan, R., and Xu, Z. (2015). Identifying effective influencers based on trust for electronic word-of-mouth marketing: a domain-aware approach. *Inform. Sci.* 306, 34–52. doi: 10.1016/j.ins.2015.01.034

Liu, X., Schuckert, M., and Law, R. (2018). Utilitarianism and knowledge growth during status seeking: evidence from text mining of online reviews. *Tour. Manag.* 66, 38–46. doi: 10.1016/j.tourman.2017.11.005

Luqiua, L., Schmierbach, M., and Nga, Y. (2019). Willingness to follow opinion leaders: a case study of Chinese Weibo. *Comput. Hum. Behav.* 101, 42–50. doi: 10.1016/j.chb.2019.07.005

Mandal, S., Shah, C., Pea-Alves, S., Hecht, M., and Greene, K. (2020). Understanding the spread of prevention and cessation messages on social media for substance use in youth. *Aslib J. Inf. Manag.* 73, 240–253. doi: 10.1108/AJIM-07-2019-0166

Martins, A. C. R. (2008). Mobility and social network effects on extremist opinions. *Phys. Rev. E* 78:036104. doi: 10.1103/PhysRevE.78.036104

Nasir, V. A., Keserel, A. C., Surgit, O. E., and Nalbant, M. (2021). Segmenting consumers based on social media advertising perceptions: how does purchase intention differ across segments? *Telem. Inform.* 64:101687. doi: 10.1016/j.tele.2021.101687

Pérez, I. J., Cabrerizo, F. J., Alonso, S., Dong, Y. C., Chiclana, F., and Herrera-Viedma, E. (2018). On dynamic consensus processes in group decision making problems. *Inf. Sci.* 459, 20–35. doi: 10.1016/j.ins.2018.05.017

Pötzschke, S., and Braun, M. (2017). Migrant sampling using Facebook advertisements: a case study of polish migrants in four European countries. *Soc. Sci. Comput. Rev.* 35, 633–653. doi: 10.1177/0894439316666262

Previte, J. (1999). Internet advertising: an assessment of consumer attitudes. *Prometheus* 17, 199–209. doi: 10.1080/08109029908629550

Priester, J. R., and Petty, R. E. (1995). Source attributions and persuasion: perceived honesty as a determinant of message scrutiny. *Personal. Soc. Psychol. Bull.* 21, 637–654. doi: 10.1177/0146167295216010

Rogers, E. (2003). *The Diffusion of Innovations. 5th Edition*. eds. Z. Rezvani, J. Jansson and Bodin (New York: Free Press).

Shmargad, Y., and Sanchez, L. (2020). Social media influence and electoral competition. *Soc. Sci. Comput. Rev.* 40, 4–23. doi: 10.1177/0894439320906803

Traud, A., Mucha, P., and Porter, M. (2012). Social structure of Facebook networks. *Physica A* 391, 4165–4180. doi: 10.1016/j.physa.2011.12.021

Wang, Z., Liu, H., Liu, W., and Wang, S. (2020). Understanding the power of opinion leaders' influence on the diffusion process of popular mobile games: travel frog on Sina Weibo. *Comput. Hum. Behav.* 109:106354. doi: 10.1016/j.chb.2020.106354

Wang, J., Wang, X., and Li, Y. (2019). A discrete electronic word-of-mouth propagation model and its application in online social networks. *Physica A* 527:121172. doi: 10.1016/j.physa.2019.121172

Wu, Z., Zhou, Q., Dong, Y., Xu, J., Altalhi, A. H., and Herrera, F. (2022). Mixed opinion dynamics based on DeGroot model and Hegselmann-Krause model in social networks. *IEEE Trans. Syst. Man Cyber. Syst.* 1–13. doi: 10.1109/TSMC.2022.3178230 (in press).

Yang, C., and Hsu, T. (2017). Effects of skepticism about corporate social responsibility advertising on consumer attitude. *Soc. Behav. Pers.* 45, 453–467. doi: 10.2224/sbp.5788

Zhan, M., Kou, G., Dong, Y., Chiclana, F., and Herrera-Viedma, E. (2022). *Bounded Confidence Evolution of Opinions and Actions in Social Networks. IEEE Transactions Cybernetics* 52, 7017–7028. doi: 10.1109/TCYB.2020.3043635

Zhang, Z., Gao, Y., and Li, Z. (2020). Consensus reaching for social network group decision making by considering leadership and bounded confidence. *Knowl.-Based Syst.* 204:106240. doi: 10.1016/j.knosys.2020.106240

Zhang, Z., Li, Z., and Gao, Y. (2021). Consensus reaching for group decision making with multi-granular unbalanced linguistic information: a bounded confidence and minimum adjustment-based approach. *Inform. Fus.* 74, 96–110. doi: 10.1016/j.inffus.2021.04.006

Zhang, L., Su, C., Jin, Y., Goh, M., and Wu, Z. (2018). Cross-network dissemination model of public opinion in coupled networks. *Inf. Sci.* 451, 240–252. doi: 10.1016/j.ins.2018.04.037

Zhang, H., Xiao, J., and Dong, Y. (2019). Integrating a consensus-reaching mechanism with bounded confidences into failure mode and effect analysis under incomplete context. *Knowl. Based Syst.* 183:104873. doi: 10.1016/j.knosys.2019.104873

Zhang, Z., Zhang, Z., and Yang, Y. (2016). The power of expert identity: how website recognized expert reviews influence travelers' online rating behavior. *Tour. Manag.* 55, 15–24. doi: 10.1016/j.tourman.2016.01.004

Zhao, Y., and Kou, G. (2014). Bounded confidence-based opinion formation for opinion leaders and opinion followers on social networks. *Stud. Inform. Control* 23, 153–162. doi: 10.24846/v23i2y201403

Zhao, Y., Kou, G., Peng, Y., and Chen, Y. (2018). Understanding influence power of opinion leaders in e-commerce networks: an opinion dynamics theory perspective. *Inf. Sci.* 426, 131–147. doi: 10.1016/j.ins.2017.10.031

Zhou, Q., Wu, Z., Altalhi, A. H., and Herrera, F. (2020). A two-step communication opinion dynamics model with self-persistence and influence index for social networks based on the degroot model. *Inf. Sci.* 519, 363–381. doi: 10.1016/j.ins.2020.01.052

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 07 December 2022
DOI 10.3389/fpsyg.2022.954132

Check for updates

OPEN ACCESS

EDITED BY
Oliver Niebuhr,
University of Southern Denmark, Denmark

REVIEWED BY
Paolo Giudici,
University of Pavia, Italy
Yuanjun Zhao,
Nanjing Audit University, China

*CORRESPONDENCE
Xin Liu
xinliu_1005@163.com

SPECIALTY SECTION
This article was submitted to Personality and Social Psychology, a section of the journal Frontiers in Psychology

RECEIVED 27 May 2022
ACCEPTED 08 November 2022
PUBLISHED 07 December 2022

CITATION
Wang Q, Liu X and Zhang C (2022) Evolutionary game analysis of FinTech transformation: A social co-governance pattern of peer-to-peer lending market in China. *Front. Psychol.* 13:954132. doi: 10.3389/fpsyg.2022.954132

COPYRIGHT
© 2022 Wang, Liu and Zhang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Evolutionary game analysis of FinTech transformation: A social co-governance pattern of peer-to-peer lending market in China

Qi Wang, Xin Liu* and Chenghu Zhang

School of Economics and Finance, Xi'an Jiaotong University, Xi'an, China

Benign exit has become the main theme of the transformation in China's peer-to-peer (P2P) lending industry. To protect the interests of investors in the benign exit process, this paper proposes a social co-governance pattern using a tripartite evolutionary game model to capture the behavior strategies of P2P lending platforms, investors, and financial regulators. The results demonstrate that there are four evolutionary stable strategies for the game model, among which the positive disposal of P2P lending platforms, the participation of the investors, and the co-governance policy of financial regulators is the optimal strategy in the benign exit process. The results also show that the initial proportion of P2P lending platforms, investors, and financial regulators would significantly affect the convergence speed of the evolutionary stable strategy. The proposed social co-governance pattern would effectively safeguard the interests of investors if incentive, penalty, and reputation mechanisms are well-designed. This paper provides in-depth implications for protecting investors' interests in the transformation of the P2P lending industry and enhancing the sustainable development of the FinTech industry.

KEYWORDS

Financial Technology (FinTech), social co-governance, tripartite evolutionary game, transformation, peer-to-peer lending

Introduction

Digital technology, such as big data, block chain, and cloud computing, is quickly evolving in the third technological revolution (Li G. et al., 2022). The game-changing technological innovations have been triggered and introduced into Financial Technology (FinTech), transforming the way the financial industry operates and fulfilling customers' needs (Agarwal and Zhang, 2020; Brandl and Hornuf, 2020; Chen and Sivakumar, 2021; Weng and Luo, 2021; Lei et al., 2022). Taking its place as one of the most significant segments in the broad area of FinTech (Luther, 2019), peer-to-peer (P2P) lending has

FIGURE 1
Geographical distribution of defaulted P2P lending platforms in China.

emerged as an infomediary platform that links up investors and borrowers to form debt–credit relationships *via* the Internet (Lee and Lee, 2012; Lin et al., 2013; An et al., 2022). Compared with the traditional financing channel, P2P lending platforms not only facilitate a convenient approach for small- and medium-sized enterprises to attain short-term loans (Feng et al., 2015) but also provide feasible investment options to investors for their idle capital (Bachmann et al., 2011; Wei and Lin, 2017). After the vigorous development of the last decade, the number of P2P lending platforms in China has reached 6,607 by 2022, with the total amount of loans exceeding 1.27 trillion dollars. Despite its merits, P2P lending has received strong criticism for its uncertain legal (Yang et al., 2018), regulatory arbitrage (Deng, 2022), and credit risks (Giudici et al., 2019). Numerous defaulted platforms occurred from time to time in the P2P lending market and may culminate in substantial financial losses, including fund-raising fraud, suspended operations, lost investment, and even business close-down (Fu et al., 2020). There is a remarkable increase in the number of defaulted P2P lending platforms in China (Yoon et al., 2019), and a growing default rate that achieves the highest point in history, i.e., 44.37% by 2021. The defaulted platforms are mainly located in the southeast and central regions of China, with the southeast region being the most severe (refer to Figure 1), which seriously infringes the interests of consumers and undermines the sustainability of the FinTech industry.

The increasing risks in the P2P lending industry have heightened the pressing need for financial regulators to develop policies to guide P2P lending platforms for transformation. In 2019, the Regulation and Rectification Office (RRO)[1] issued the regulatory policy "Guidance on P2P lending platform disposition by classification and risk prevention," urging P2P lending platforms to follow the guideline of benign exit.[2] However, from a realistic point of view, the benign exit process of P2P lending platforms is not satisfactory concerning the protection of investors' interests. Hundreds of large-scale P2P lending platforms such as "*Tuandaiwang*," "*Wanglibao*," and "*Aiqianjin*," which have been investigated by The Ministry of Public Security of China, resulted in billions of dollars in property losses for investors. Hence, there is an urgent need to address these problems by building up a benign exit pattern in the P2P lending industry to protect the legitimate rights and interests of investors, with respect to providing in-depth insights for the future sustainable development of the FinTech industry.

Since the P2P lending industry has shown irrational prosperity in the last decade due to the long-term inclusive policies in China (He and Li, 2021), thus the benign exit of P2P lending platforms is inseparable from the rigorous policy of financial regulators. However, the primary issue of the P2P lending platforms facing during the process of benign exit is the huge credit assignment with investors, which would be difficult to solve either by financial regulators or by investors. Thus, the co-governance of financial regulators, investors, and P2P lending platforms in the benign exist process can meet the interests of all stakeholders, which are the embodiment of corporate social responsibility (Jo and Harjoto, 2012; Kim et al., 2018; Wang et al., 2020). Compared with traditional regulation, the co-governance argues that the government, the market, and related stakeholders should be in equal positions to form a coordinated and effective network, in order to effectively distribute social benefits and ensure the maximization of social interests (Ackerman, 2004; Wu et al., 2018). Co-governance is more proactive and creative in the benign exit process of P2P lending platforms, which can drive investors to prevent P2P lending platforms from vicious exit through social supervision and encourage P2P lending platforms to comply with the law through industry self-regulation and market incentives. In addition, co-governance can improve the flexibility of regulation, broaden the applicability of policies, and reduce the cost of the benign exit process.

From the co-governance perspective, this study provides novel insights into the benign exit of P2P lending platforms and has three main contributions to filling gaps in the existing literature. First, the benign exit pattern of P2P lending platforms will be discussed from the perspective of co-governance, which can provide deeper implications for the research of defaulted P2P lending platforms. Second, we construct a tripartite evolutionary game to capture the behavioral strategies of P2P lending platforms, investors, and financial regulators, to develop a more effective co-governance pattern of the benign exist process. Third, the evolutionary stable strategy for enabling the benign exit of the P2P lending platform is determined. A

1 The Regulation and Rectification Office (RRO) is affiliated with China Banking and Insurance Regulatory Commission, which is primarily responsible for preventing and rectifying the risks of Internet finance and P2P lending industry.

2 See http://www.nifd.cn/ResearchComment/Details/1229

sensitivity analysis of the influencing factors is performed to understand the effective conditions that can drive the tripartite evolutionary game model to its ideal status faster. This study provides policy implications by shedding light on the premise of interaction among P2P lending platforms, investors, and financial regulators and illustrates how to weigh strategies of the stakeholders to maximize payoffs and achieve a stable benign exit of P2P lending platforms.

The remainder of this paper is organized as below. Section Literature review reviews the literature. Section Problem description and basic assumptions describes the problem and establishes a tripartite evolutionary game model including P2P lending platforms, investors, and financial regulators. Section Stability analysis conducts numerical simulations to verify the theoretical results and sensitivity analysis to provide deeper insights into the studied problem. Section Numerical simulation presents a discussion of the results. Section Discussion summarizes the conclusions and puts forward relevant policy implications. Section Conclusion draws the conclusion.

Literature review

This research is closely related to three streams of literature: the risk of P2P lending platforms and their regulation, the co-governance and its applications, and the application of evolutionary games in the research of P2P lending platforms.

Defaulted platform and its regulation

Peer-to-peer lending platforms connect investors and borrowers *via* qualified third-party internet platforms, which not only facilitate a convenient approach for small and medium enterprises to get short-term loans but also provide feasible investment options to investors for their idle capital (Bachmann et al., 2011). However, with the main technological drivers, i.e., big data analytics, artificial intelligence, and block chain technology, risks arise with the development of the most important financial technologies, i.e., peer-to-peer lending (Giudici, 2018; Aldasoro et al., 2022). Particularly, the P2P lending platforms in China are prone to default due to the absence of stringent market regulations (Zhang and Wang, 2019) and low entry barriers (Chen and Tsai, 2017), causing serious damage to the interests of consumers and the stability of the industry, which has gradually aroused scholars' concerns (Emekter et al., 2015; Li et al., 2016; Yang and Luo, 2017; Liu et al., 2018; Xia et al., 2019). Prior studies mainly focused on the identification of the key factors that affect platform risk and the application of models to predict risk to mitigate the platform risks in P2P lending industry (Giudici et al., 2020). The impacts of platform characteristics and macro-financial environment factors on the default risk of P2P lending platforms are examined (Yoon et al., 2019). It is demonstrated that P2P lending platforms that closed down as a result of liquidity issues were usually due to a lack of high-quality risk management techniques (Liu et al., 2019). From the investors' side, it is found that the higher the level of risk aversion of the investors, the higher the level of risks of the P2P lending platform (Yan et al., 2018; Cheng and Guo, 2020). By digging deep into the very nature of the platforms, it is also shown that the impact of network effects in peer-to-peer lending platforms would increase the opportunities and risks for both investors and borrowers (Chen et al., 2022).

A large and growing body of literature has investigated the regulation of defaulted P2P lending platforms. It is argued that P2P lending is an example of how modern technology enables the integration of a range of economic functions, therefore, a new approach to market regulation is warranted which is more consistent with emerging institutional arrangements (Davis, 2016). It is pointed out that the regulation of P2P lending should be consistent with leveraged information mechanisms to reduce information asymmetry and market friction and ensure market transparency, competition, and fair pricing (Yang et al., 2018). It is also pointed out that the financial regulators should first motivate and then regulate the P2P lending platforms (Zhang and Wang, 2019). In comparison to the regulatory regime of P2P lending markets in the US, the UK, and Japan, the extent to which the new regulatory regime was likely to reduce the default of P2P lending platforms to protect the interests of the investors is examined (Huang, 2018). It is concluded that a "national bank charter" was the best way to ensure greater regulation of the P2P lending industry, provide sufficient incentives for investors, reduce systemic risk, and allow for greater regulatory knowledge of related institutions to ensure their compliance of consumer protection laws (Luther, 2019).

Co-governance and its applications

In the late twentieth century, the high welfare policies of developed countries resulted in overstaffed and inefficient government agencies (Offe, 1984). It is noted that the government, the market, and the social actors are supposed to be in equal positions and form a coordinated and effective network in order to more effectively distribute and ensure the maximization of social benefits (Gelatt, 1992). On this basis, an inclusive and flexible concept of co-governance was formed. The co-governance was defined as an approach in which a mixture of instruments is brought to bear on a specific problem, emphasizing the coordination between public and private agents in the regulatory process (Eijlander, 2005). According to a prior study (Bartle and Vass, 2005), co-governance may arise in the process of creating regulatory rules by incorporating the opinions of the government, non-governmental organizations, market players, individuals, and

other stakeholders. Co-governance in enforcement involves all modes of regulation in which regulations are designed and set by public authorities and enforced by the coordinated actions of public authorities and regulated firms (Rouvière and Ca Swell, 2012). It is concluded that the degree of cooperation and competition depends on the existing regulatory arrangements, the congruence of goals of the different actor groups, and the institutionalization of industrial relations (Tosun et al., 2016).

Researchers have explored these potential complementarities and gains from coordination in many different areas of regulation. In the environmental protection landscape, the legal and institutional frameworks of environmental co-governance were constructed to adopt co-construction, co-governance, and the sharing of innovative social governance patterns (Birnbaum, 2016; Iaione, 2016; Ko et al., 2019; Abdullah et al., 2020; Jin et al., 2020; Xu et al., 2020). In more recent years, the concept of co-governance has gradually aroused policy makers' attention in the field of food safety regulation, arguing that co-governance is a kind of societal-wide innovation that integrates diverse resources and efforts from multiple stakeholders including government, industry, and social forces for better and sustainable development of an economy's food institution and system (Martinez et al., 2007; Wu et al., 2018; Chen and Li, 2019; Chen and Wu, 2019; Meng et al., 2021; Pan et al., 2021; Yan et al., 2021). It is found that a significant degree of cooperation between private regulators and public supervisors was the key to ensuring the effectiveness of regulation and pointed out that public supervision and enforcement must be responsive to the peculiarities of co-governance arrangements (Cherednychenko, 2016). Few works of literature pay attention to the application of co-governance in the E-commerce field. Through an evolutionary game model among the government, E-commerce platforms, and rights holders, it is found that reasonable adjustment of the reward and punishment measures of government supervisory agencies can produce positive guidance to platform and operators, and the related social environment, social benefits, the sense of acquisition by the government, platforms and rights holders can be strengthened (Li J. et al., 2022).

Application of evolutionary games in the P2P lending market

Compared to classic game theory, evolutionary game theory has the merits of considering the bounded rationality of the players with the ability to keep learning, adapt to the market environment, and adjust their strategies, which can describe the dynamic process of decision-making theoretically (Taylor, 1979; Smith, 1988). Particularly, evolutionary game theory was widely used to allow a deeper insight into the bounded rationality of P2P lending platforms, investors, regulators, and other stakeholders. Previous research has established evolutionary game models among regulatory authorities, P2P lending platforms, and borrowers to evaluate the strict supervision strategy of the P2P lending platform (You et al., 2021) and determine certain conditions when the three players will converge on the strategies of positive supervision, self-discipline operation, and compliance, respectively (Chunsheng, 2020). An evolutionary game model was established between local governments and P2P lending platforms regarding the benign exit of P2P lending platforms in China, and it is noted that the two players can converge to the optimal equilibrium of "incentives and benign exits" under certain conditions (Zhang et al., 2020). An evolutionary game model of different regulatory stages was established to analyze the optimal state of the P2P lending market structure under the strengthening supervision and shrinking industry (Peng et al., 2020). Evolutionary game models were also developed in several lines of research to help understand the risk supervision of P2P lending platforms (Gu et al., 2018), to investigate the influence pathway of the guarantee mechanism on users' participation (Weng and Luo, 2021), and to analyze the risk preference behavior of lenders and the credit choice of borrowers (Liu and Xia, 2017).

The literature presented, thus, far provides sufficient studies on the default risk regulation of P2P lending platforms. Undoubtedly, these studies have laid a solid foundation for understanding the default risk of P2P lending platforms and provide practical implications to policy makers. However, few scholars have been able to draw on any rigorous research into the benign exit process of P2P lending platforms, and the crucial role of investors in the regulation of P2P lending platforms has been ignored. The study would have been more useful if a co-governance pattern in the benign exit process of P2P lending platforms in China was considered. To address this problem, this paper established a tripartite evolutionary game model including P2P lending platforms, investors, and financial regulators, studied the formulation of the co-governance pattern, concentrated on the analysis of the influencing factors of the co-governance pattern on the evolution stability strategy (ESS; Smith and Price, 1973).

Problem description and basic assumptions

Problem description

Based on a co-governance framework, the P2P lending platforms, investors, and financial regulators are the main players in the benign exit process of P2P lending platforms. P2P lending platforms need to comply with the benign exit rules, address their debt risks by returning credit assignments, and meet the demands of investors and financial regulators.

Moreover, the P2P lending platforms are supposed to protect the legitimate rights and interests of investors to the maximum extent in a timely manner.

The investors, however, are not only the direct stakeholders of the benign exit of P2P lending platforms but also the direct beneficiaries of the regulation of P2P lending platforms. Hence, it is essential for investors to participate in the co-governance of the benign exit process. Specifically, the investors' supervision of P2P lending platforms and the performance assessment of financial regulators would play a significant role in their strategies.

There is strong evidence that financial regulators play a crucial role in regulating P2P lending platforms. The financial regulators, which aim to accelerate the benign exit process and protect the rights and interests of investors, undertake multiple tasks such as imposing a penalty on P2P lending platforms that do not comply with the benign exit rules and rewarding investors' positive engagement in co-governance. Effective financial regulation will have a conductive role in restraining the operations strategy of P2P lending platforms and protecting the interests of investors. Otherwise, it will jeopardize the public assessment of financial regulators' performance. Above all, the co-governance pattern of benign exit of P2P lending platforms is established based on the interaction of P2P lending platforms, investors, and financial regulators, as shown in Figure 2.

In the co-governance pattern, P2P lending platforms, investors, and financial regulators are supposed to coordinate and cooperate in building principles, ensuring process implementation, enforcement, and supervision to accelerate the benign exit process. Therefore, the repayment ratio of the non-performing assets will be increased, and the asset degradation will be reduced. Specifically, building principles refer to verifying the asset and capital of P2P lending platforms and forming benign exit principles. Process implementation refers to putting these principles into real practice. Enforcement refers to ensuring the benign exit of P2P lending platforms in compliance with laws and regulatory policies. Supervision refers to the ongoing investigation of the whole process of benign exit of P2P lending platforms.

Payoffs of three participants

Scenario with co-governance regulation

This sub-section will discuss the expected payoffs of P2P lending platforms, investors, and financial regulators based on the co-governance pattern.

P2P lending platforms' expected payoffs

When P2P lending platforms adopt the "positive disposal" strategy and investors adopt the "participating in co-governance" strategy, then the P2P lending platforms will receive the reputation benefits from the investors. Thus, their expected payoffs are equal to $-C_{P1} + \psi_{P1} - \theta_1 pD(1 + i)$. If investors adopt the "not participating in co-governance" strategy, then the P2P lending platforms cannot gain the reputation benefits from investors, and the expected payoffs are equal to $-C_{P1} - \theta_1 pD(1 + i)$. When P2P lending platforms adopt the "negative disposal" strategy, their costs will be reduced from C_{P1} to C_{P2}; however, once this behavior is found by the financial regulators, they will suffer a penalty, in this case, their expected payoffs in case investors adopt "participating in co-governance" or "not participating in co-governance" are equal to $-C_{P2} - \alpha F - \alpha\theta_2 pD(1 + i) - \psi_{P2}$ and $-C_{P2} - \beta F - \beta\theta_2 pD(1 + i)$, respectively.

Investors' expected payoffs

Investors can obtain a fraction of credit assignment from the P2P lending platforms plus the rewards from the financial regulators when they adopt the "participating in co-governance" strategy, thus, their expected payoffs in case P2P lending platforms adopt "positive disposal" and "negative disposal" strategy are equal to $-(1-\theta_1)pD(1+i)+R_L-C_L/D$ and $-\alpha(1-\theta_2)pD(1 + i) + R_L - C_L/D$, respectively. If the investors adopt the "not participating in co-governance" strategy, they will not gain the rewards from the financial regulators; in this case, their payoffs will be dependent on the P2P lending platforms' strategy choice and the regulatory efficiency of the financial regulators, thus, the expected payoffs of investors in case P2P lending platforms adopt "positive disposal" and "negative disposal" strategy will be equal to $-(1-\theta_1)pD(1+i)$ and $-\beta(1-\theta_2)pD(1+i)$, respectively.

Financial regulators' expected payoff

When the financial regulators adopt the "co-governance regulation" strategy, they will incur an initiative cost to set up the social co-governance pattern. In this case, when P2P lending platforms adopt a "positive disposal" strategy and investors adopt a "co-governance" strategy, then the expected payoffs of the financial regulators are equal to $\psi_{R1} + \psi_{W1} - M_R - C_{R1} - R_L$. If investors adopt the "not participating in co-governance" strategy, then the expected payoffs of the financial regulators are equal to $\psi_{W1} - M_R - C_{R1}$. When P2P lending platforms adopt a "negative disposal" strategy, there may be a social welfare loss in accordance with the regulation efficiency. In this case, if investors adopt a "participating in co-governance" strategy, then the financial regulators will be required to pay rewards to the investors; in this case, the expected payoffs of financial regulators are equal to $\psi_{R1}+\alpha F-M_R-C_{R1}-R_L-(1-\alpha)\psi_{W2}$. If investors adopt the "not participating in co-governance" strategy, the payoffs of financial regulators will be equal to $\beta F - M_R - C_{R1} - (1 - \beta)\psi_{W2}$. Table 1 shows the payoff matrix of three participants with co-governance regulation.

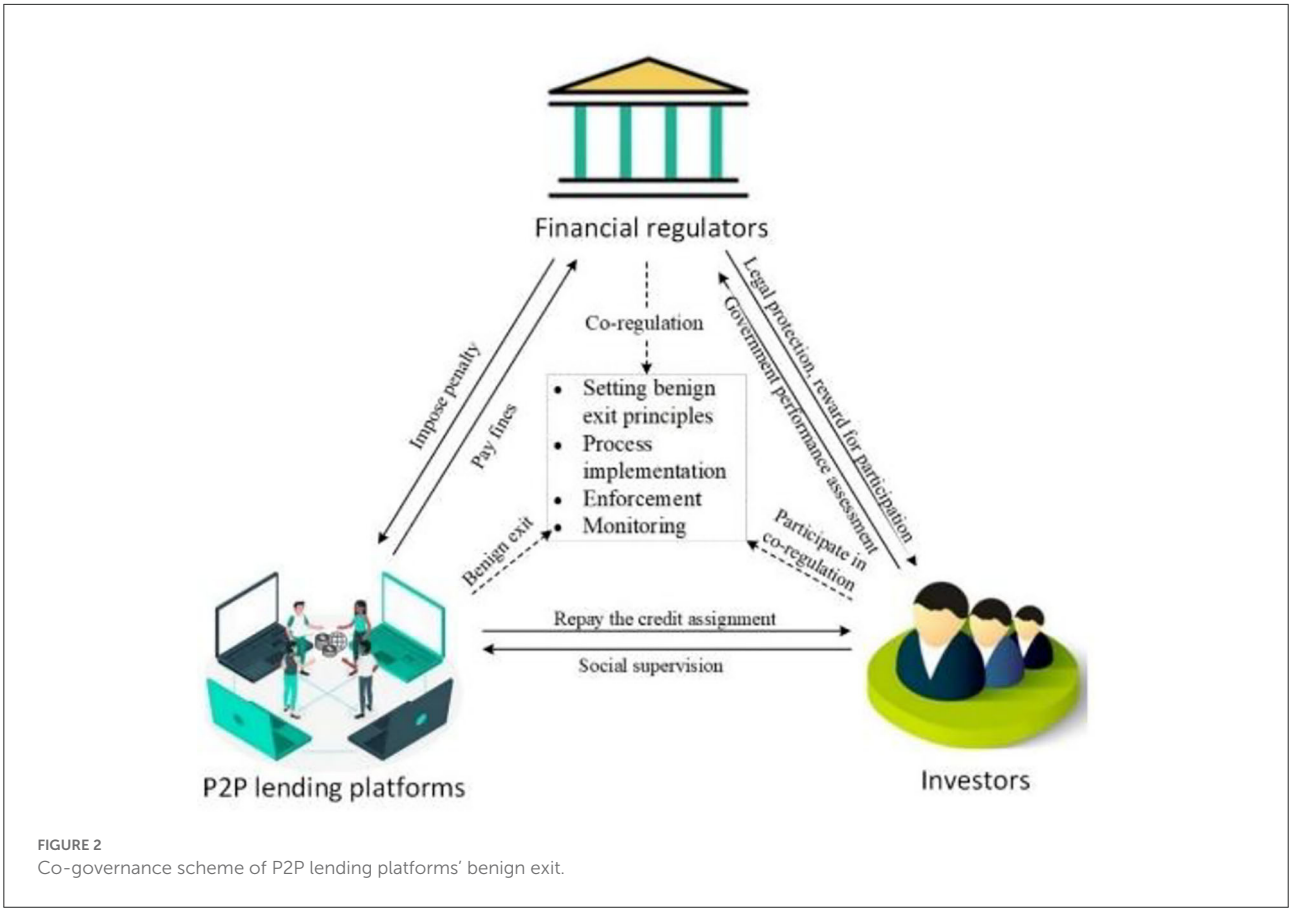


FIGURE 2
Co-governance scheme of P2P lending platforms' benign exit.

TABLE 1 The payoff matrix of the tripartite game.

P2P lending platforms	Investors	Financial regulators	
		Co-governance (z)	Traditional regulation (1−z)
Positive exit (x)	Participating in co-governance (y)	$(\alpha_1,\ \beta_1,\ \gamma_1)$	$(\alpha_2,\ \beta_2,\ \gamma_2)$
	Not participating in co- regulation $(1-y)$	$(\alpha_3,\ \beta_3,\ \gamma_3)$	$(\alpha_4,\ \beta_4,\ \gamma_4)$
Negative exit $(1-x)$	Participating in co- regulation (y)	$(\alpha_5,\ \beta_5,\ \gamma_5)$	$(\alpha_6,\ \beta_6,\ \gamma_6)$
	Not participating in co- regulation $(1-y)$	$(\alpha_7,\ \beta_7,\ \gamma_7)$	$(\alpha_8,\ \beta_8,\ \gamma_8)$

Scenario with traditional regulation

This sub-section will discuss the expected payoffs of P2P lending platforms, investors, and financial regulators in the scenario with traditional regulation.

P2P lending platforms' expected payoffs

When P2P lending platforms adopt the "positive disposal" strategy, their expected payoffs are the same as in the scenario with co-governance regulation. It should be noted that there is no direct causal relationship between the regulation type and P2P lending platforms' reputation benefits, their reputation benefits are associated with the investors' strategy choice. When P2P lending platforms adopt a "negative regulation" strategy, they will suffer a penalty from the financial regulators in accordance with the regulation efficiency, thus, their expected payoffs in case investors adopt "participating in co-governance" or "not participating in co-governance" are equal to $-C_{P2}-\alpha F-\alpha\theta_2 pD(1+i)-\psi_{P2}$ and $-C_{P2}-\beta F-\beta\theta_2 pD(1+i)$, respectively.

Investors' expected payoffs

Compared with the scenario with co-governance, whether the investors adopt the "participating in co-governance" strategy or "not participating in co-governance" strategy, they will not gain a reward from the financial regulators, thus, when the investors adopt the "participating in co-governance" strategy, the expected payoffs of investors in case the P2P lending

TABLE 2 The specific payoff of P2P lending platforms, investors, and financial regulators.

Payoff	P2P lending platforms	Investors	Financial regulators
$(\alpha_1,\ \beta_1,\ \gamma_1)$	$-C_{P1}+\psi_{P1}-\theta_1 pD(1+i)$	$-(1-\theta_1)pD(1+i)+R_L-C_L$	$\psi_{R1}+\psi_{W1}-M_R-C_{R1}-R_L$
$(\alpha_2,\ \beta_2,\ \gamma_2)$	$-C_{P1}+\psi_{P1}-\theta_1 pD(1+i)$	$-(1-\theta_1)pD(1+i)-C_L$	$-C_{R2}+\psi_{W1}-\psi_{R2}$
$(\alpha_3,\ \beta_3,\ \gamma_3)$	$-C_{P1}-\theta_1 pD(1+i)$	$-(1-\theta_1)pD(1+i)$	$\psi_{W1}-M_R-C_{R1}$,
$(\alpha_4,\ \beta_4,\ \gamma_4)$	$-C_{P1}-\theta_1 pD(1+i)$	$-(1-\theta_1)pD(1+i)$	$\psi_{W1}-C_{R2}$
$(\alpha_5,\ \beta_5,\ \gamma_5)$	$-C_{P2}-\alpha F-\alpha\theta_2 pD(1+i)-\psi_{P2}$	$-\alpha(1-\theta_2)pD(1+i)+R_L-C_L$	$\psi_{R1}+\alpha F-M_R-C_{R1}-R_L-(1-\alpha)\psi_{W2}$
$(\alpha_6,\ \beta_6,\ \gamma_6)$	$-C_{P2}-\psi_{P2}$	$-pD(1+i)-C_L$	$-C_{R2}-\psi_{W2}-\psi_{R2}$
$(\alpha_7,\ \beta_7,\ \gamma_7)$	$-C_{P2}-\beta F-\beta\theta_2 pD(1+i)$	$-\beta(1-\theta_2)pD(1+i)$	$\beta F-M_R-C_{R1}-(1-\beta)\psi_{W2}$
$(\alpha_8,\ \beta_8,\ \gamma_8)$	$-C_{P2}$	$-pD(1+i)$	$-C_{R2}-\psi_{W2}$

platforms adopt "positive disposal" or "negative disposal" strategy are equal to $-(1-\theta_1)pD(1+i)-C_L/D$ and $-pD(1+i)-C_L/D$, respectively. When investors adopt the "not participating in co-governance" strategy, the expected payoffs of investors in case the P2P lending platforms adopt a "positive disposal" or "negative disposal" strategy are equal to $-(1-\theta_1)pD(1+i)$ and $-pD(1+i)$, respectively.

Financial regulators' expected payoffs

When P2P lending platforms adopt the "positive disposal" strategy, the financial regulators will gain a social welfare improvement, thus, the expected payoffs of the financial regulators, in case investors adopt the "participating in co-governance" or "not participating in co-governance" strategy, are equal to $-C_{R2}+\psi_{W1}-\psi_{R2}$ and $\psi_{W1}-C_{R2}$, respectively. Moreover, when P2P lending platforms adopt the "negative disposal" strategy, the financial regulators will not find the P2P lending platform's "negative disposal" behavior in this scenario, thus, they will suffer a social welfare loss, then the expected payoffs of the financial regulators in case investors adopt "participating in co-governance" or "not participating in co-governance" strategy, are equal to $-C_{R2}-\psi_{W2}-\psi_{R2}$ and $-C_{R2}-\psi_{W2}$, respectively. Table 2 presents the specific payoff matrix of the three participants.

Stability analysis

Basic assumptions and model parameters

To analyze the defined problem, the main assumptions are summarized as follows:

Assumption 1: The participants in the game process include P2P lending platforms, investors, and financial regulators, as well as all of whom are bounded rational.

Assumption 2: The strategy set of P2P lending platforms is {positive disposal, negative disposal}, the strategy set of Investors is {participating in co-governance, not participating in co-governance}, and the strategy set of financial regulators is {co-governance regulation, traditional regulation}.

Assumption 3: The proportion of individuals in P2P lending platforms that adopt "positive disposal" strategy is x, and the proportion of individuals in P2P lending platforms that adopt "negative disposal" strategy is $1-x$. The proportion of individuals in investors that adopt "participating in co-governance" strategy is y, and the proportion of individuals in investors that adopts "not participating in co-governance" strategy is $1-y$. The proportion of individuals in financial regulators that adopt "co-governance regulation" strategy is z, and the proportion of individuals in financial regulators that adopt "traditional regulation" strategy is $1-z$.

Assumption 4: During the process of the game, every player in the game wants to get their maximum benefits. However, they cannot get their optimal strategies in a game due to their bounded rationality. They can change their strategies through learning and observing in the continuous games until achieving their stable states.

Assumption 5: When P2P lending platforms adopt the "negative disposal" strategy, they will pay less costs. However, once the "negative disposal" behavior is investigated by the financial regulators, the financial regulators will impose penalty on P2P lending platforms to push them to adopt the "positive disposal" strategy.

The symbols and connotations of the parameters used in the model are listed in Table 3.

Replicated dynamic equation

Set the expected payoff of P2P lending institutions' "positive disposal" strategy as U_{P1}, the expected payoff of P2P lending institutions' "negative disposal" strategy as U_{P2}, the average expected payoff of P2P lending institutions as $\overline{U}_P$. The expected payoff of investors' "participating in co-governance" strategy as U_{L1}, the expected payoff of investors' "not participating in co-governance" strategy as U_{L2}, the average expected payoff of investors as $\overline{U}_L$. The expected payoff of financial regulators' "co-governance regulation" strategy as U_{R1}, the expected payoff of financial regulators' "traditional regulation" strategy

TABLE 3 Symbols and notations of the parameters in the model.

Parameters	Connotations
C_{P1}	The cost of P2P lending platforms when adopting "positive disposal" strategy.
C_{P2}	The cost of P2P lending platforms when adopting "negative disposal" strategy.
θ_1	The proportion of credit assignment to the investors when P2P lending platforms adopting "positive disposal" strategy.
θ_2	The proportion of creditor's rights transferred to the investors when P2P lending platforms adopting "negative disposal" strategy.
F	The penalty of financial regulators to P2P lending platforms when P2P lending platforms adopt "negative disposal" strategy.
ψ_{P1}	The reputation benefits of P2P lending platforms.
ψ_{P2}	The reputation loss of P2P lending platforms.
R_L	The rewards of financial regulators to investors when investors adopt "participating in co-governance" strategy.
C_L	The cost of investors when adopting "participating in co-governance" strategy.
D	The centrality degree of investors.
p	The investment volume of investors.
i	The interest rates.
C_{R1}	The cost of financial regulators when adopting "co-governance regulation."
C_{R2}	The cost of financial regulators when adopting "traditional regulation" strategy, $C_{R2} < C_{R1}$.
α	The regulatory efficiency of co-governance regulation.
β	The regulatory efficiency of traditional regulation.
ψ_{w1}	The social welfare improvement when P2P lending platforms adopt "positive disposal" strategy.
ψ_{w2}	The social welfare loss when P2P lending platforms adopt "positive disposal" strategy.
ψ_{R1}	The reputation benefits of financial regulators when adopting "co-governance regulation" strategy.
ψ_{R2}	The reputation loss of financial regulators when adopting "traditional regulation" strategy.
M_R	The initiative cost of financial regulators to set up the co-governance regulation pattern.

as U_{R2}, the average expected payoff of financial regulators as $\overline{U}_R$.

According to Table 2, the expected payoffs that P2P lending platforms gain when they adopt the "positive disposal" strategy is as follows:

$$\begin{aligned} U_{P1} &= yz[-C_{P1} + \psi_{P1} - \theta_1 pD(1+i)] \\ &+ y(1-z)[-C_{P1} + \psi_{P1} - \theta_1 pD(1+i)] + (1-y)z[-C_{P1} \\ &- \theta_1 pD(1+i)] + (1-y)(1-z)[-C_{P1} - \theta_1 pD(1+i)] \quad (1) \end{aligned}$$

The expected payoffs that P2P lending institutions gain when they choose the "negative disposal" strategy is as follows:

$$\begin{aligned} U_{P2} &= yz[-C_{P2} - \alpha F - \alpha\theta_2 pD(1+i) - \psi_{P2}] \\ &+ y(1-z)(-C_{P2} - \psi_{P2}) + (1y)z[-C_{P2} - \beta F \\ &- \beta\theta_2 pD(1+i)] + (1-y)(1-z)(-C_{P2}) \quad (2) \end{aligned}$$

Accordingly, the average expected payoff of P2P lending platforms is as follows:

$$\overline{U}_P = x\Pi_{P1} + (1-x)\Pi_{P2} \quad (3)$$

Taking the proportion of the "positive disposal" strategy as an example, the replicated dynamic equation of P2P lending platforms can be expressed as follows:

$$\begin{aligned} F(x) &= \frac{dx}{dt = x(} U_{P1} - \overline{U}_P) \\ &= x(1-x)\{(\psi_{P1} + \psi_{P2})y + [\beta F + \beta\theta_2 pD(1+i)]z \\ &+ (\alpha - \beta)[F + \theta_2 pD(1+i)]yz + C_{P2} - C_{P1} \\ &- \theta_1 pD(1+i)\} \quad (4) \end{aligned}$$

The expected payoffs that investors gain when they adopt the "participating in co-governance" strategy is as follows:

$$\begin{aligned} U_{L1} &= xz[-(1-\theta_1)pD(1+i) + R_L - C_L/D] \\ &+ x(1-z)[-(1-\theta_1)pD(1+i) - C_L/D] \\ &+ (1-x)z[-\alpha(1-\theta_2)pD(1+i) + R_L - C_L/D] \\ &+ (1-x)(1-z)[-pD(1+i) - C_L/D] \quad (5) \end{aligned}$$

The expected payoffs that investors gain when they adopt the "not participating in co-governance" strategy is as follows:

$$\begin{aligned} U_{L2} &= xz[-(1-\theta_1)pD(1+i)] + x(1-z)[-(1-\theta_1)pD(1+i)] \\ &+ (1-x)z[-\beta(1-\theta_2)pD(1+i)] \\ &+ (1-x)(1-z)[-pD(1+i)] \quad (6) \end{aligned}$$

Similarly, the average expected payoff of investors is as follows:

$$\overline{U}_L = yU_{L1} + (1-y)U_{L2} \quad (7)$$

Accordingly, the replicated dynamic equation of investors adopting the "participating in co-governance" strategy can be expressed as follows:

$$\begin{aligned} F(y) &= \frac{dy}{dt = y(} U_{L1} - \overline{U}_L) = y(1-y)\{[R_L - (\alpha - \beta)(1 \\ &- \theta_2)pD(1+i)]z + (\alpha - \beta)(1-\theta_2)pD(1+i)xz \\ &- C_L/D\} \quad (8) \end{aligned}$$

Additionally, the expected payoff that financial regulators gain when they adopt the "co-governance regulation" strategy is as follows:

$$\begin{aligned} U_{R1} &= xy\left(\psi_{R1}+\psi_{W1}-M_R-C_{R1}-R_L\right) \\ &+ x\left(1-y\right)\left(\psi_{W1}-M_R-C_{R1}\right) \\ &+ (1-x)y[\psi_{R1}+\alpha F-M_R-C_{R1}-R_L-(1-\alpha)\psi_{W2}] \\ &+ (1-x)(1-y)[\beta F-M_R-C_{R1}-(1-\beta)\psi_{W2}] \end{aligned} \quad (9)$$

The expected payoff that financial regulators gain when they adopt the "traditional regulation" strategy is as follows:

$$\begin{aligned} U_{R2} &= xy(-C_{R2}+\psi_{W1}-\psi_{R2})+x(1-y)(\psi_{W1}-C_{R2}) \\ &+ (1-x)y(-C_{R2}-\psi_{W2}-\psi_{R2}) \\ &+ (1-x)(1-y)(-C_{R2}-\psi_{W2}) \end{aligned} \quad (10)$$

The average expected payoff of the financial regulators is as follows:

$$\overline{U}_R = zU_{R1}+(1-z)U_{R2} \quad (11)$$

Similarly, the following replicated dynamic equation can be obtained when the financial regulators adopt a "co-governance regulation" strategy:

$$\begin{aligned} \frac{F(z)=dz}{dt} &= z\left(U_{R1}-\overline{U}_R\right) = z(1-z)\{-\beta\left(F+\psi_{W2}\right)x \\ &+ \left[\psi_{R1}+\psi_{R2}+\left(\alpha-\beta\right)\left(F+\psi_{W2}\right)-R_L\right]y \\ &- [(\alpha-\beta)(F+\psi_{W2})]xy+C_{R2}+\beta(F+\psi_{W2}) \\ &- C_{R1}-M_R \end{aligned} \quad (12)$$

Due to the limited rationality of P2P lending platforms, investors, and financial regulators, it is difficult for them to make the best choice in a game. Therefore, Equations 4, 8, 12 can be considered as indicating an evolutionary process, forming a tripartite replicated dynamic system. Maybe, with the development of iterations, P2P lending platforms, investors, and financial regulators can find strategies to maximize their benefits, and eventually, develop an evolutionary stable strategy (ESS).

Evolutionary stable strategies

As mentioned in Section Basic assumptions and model parameters, the whole game is constantly evolving, hence, the proportions of any strategies chosen by P2P lending platforms, investors, and financial regulators are time-dependent and can be expressed as $x(t)$, $y(t)$, and $z(t) \in [0,\ 1]$, respectively. Thus, the solution of the replicated dynamic system, consisting of Equations 4, 8, 12, is as follows: $[0,\ 1] \times [0,\ 1] \times [0,\ 1]$.

Obviously, when all the dynamic equations equal 0, which means that the whole dynamic system will tend to be stable, the P2P lending platforms, investors, and financial regulators have been able to select the optimal strategy. Thus, the equilibrium points of the tripartite game model can be calculated in the following way:

$$\begin{aligned} F(x) &= \frac{dx}{dt=x(}U_{P1}-\overline{U}_P) = x(1-x)\{(\psi_{P1}+\psi_{P2})y \\ &+ [\beta F+\beta\theta_2 pD(1+i)]z+(\alpha-\beta)[F \\ &+ \theta_2 pD(1+i)]yz+C_{P2}-C_{P1}-\theta_1 pD(1+i)\} \\ &= 0 \end{aligned} \quad (13-1)$$

$$\begin{aligned} F(y) &= \frac{dy}{dt=y(}U_{L1}-\overline{U}_L) \\ &= y(1-y)\{\left[R_L-(\alpha-\beta)\left(1-\theta_2\right)pD\left(1+i\right)\right]z \\ &+ (\alpha-\beta)(1-\theta_2)pD(1+i)xz-C_L/D\} \\ &= 0 \end{aligned} \quad (13-2)$$

$$\begin{aligned} \frac{F\left(z\right)=dz}{dt} &= z\left(U_{R1}-\overline{U}_R\right) = z(1-z)\{-\beta(F+\psi_{W2})x \\ &+ [\psi_{R1}+\psi_{R2}+(\alpha-\beta)(F+\psi_{W2})-R_L]y \\ &- [(\alpha-\beta)(F+\psi_{W2})]xy+C_{R2}+\beta(F+\psi_{W2}) \\ &- C_{R1}-M_R\} = 0 \end{aligned} \quad (13-3)$$

The equilibrium points can be identified easily by solving Equations 13-1, 13-2, and 13-3. Among these, there are eight special equilibrium points: $E_1(0,0,0)$,$E_2(1,0,0)$,$E_3(0,1,0)$,$E_4(0,0,1)$,$E_5(1,0,1)$,$E_6(1,1,0)$, $E_7(0,1,1)$, and $E_8(1,1,1)$. All stakeholders adopt a pure strategy in each of these equilibrium points, which constitute the boundary of the domain. In addition, there may exist other mixed strategy equilibrium points by solving Equation 14.

$$\begin{cases} \left(\psi_{P1}+\psi_{P2}\right)y+\left[\beta F+\beta\theta_2 pD\left(1+i\right)\right]z \\ \quad +\left(\alpha-\beta\right)\left[F+\theta_2 pD\left(1+i\right)\right]yz+C_{P2}-C_{P1} \\ \quad -\theta_1 pD\left(1+i\right)=0 \\ \left[R_L-\left(\alpha-\beta\right)\left(1-\theta_2\right)pD\left(1+i\right)\right]z \\ \quad +\left(\alpha-\beta\right)\left(1-\theta_2\right)pD\left(1+i\right)xz-\frac{C_L}{D}=0 \\ -\beta(F+\psi_{W2})x+[\psi_{R1}+\psi_{R2}+(\alpha-\beta)(F+\psi_{W2})-R_L]y \\ -[(\alpha-\beta)(F+\psi_{W2})]xy+C_{R2}+\beta(F+\psi_{W2})-C_{R1} \\ \quad -M_R=0 \end{cases} \quad (14)$$

However, according to Friedman's evolutionary game theory (Friedman, 1998), only if the equilibrium points simultaneously satisfy both a strict Nash equilibrium and a pure strategy Nash equilibrium, then they will turn into an asymptotically stable equilibrium, and thus the equilibrium point will be an ESS. Consequently, we only need to analyze the stability of the eight pure strategy Nash equilibrium points.

Moreover, according to Wainwright (1989) and Lyapunov (1992), the asymptotic stability at the equilibrium point can be evaluated by analyzing the eigenvalues of the Jacobian matrices of the system, whereby the necessary and sufficient condition for the asymptotic stability of the system is that all the eigenvalues of the Jacobian matrix are negative. Thus, the Jacobian matrix of

the tripartite dynamic game can be calculated as follows:

$$J = \begin{pmatrix} \frac{\partial F(x)}{\partial x} & \frac{\partial F(x)}{\partial y} & \frac{\partial F(x)}{\partial z} \\ \frac{\partial F(y)}{\partial x} & \frac{\partial F(y)}{\partial y} & \frac{\partial F(y)}{\partial z} \\ \frac{\partial F(z)}{\partial x} & \frac{\partial F(z)}{\partial y} & \frac{\partial F(z)}{\partial z} \end{pmatrix} = \begin{pmatrix} J_{11} & J_{12} & J_{13} \\ J_{21} & J_{22} & J_{23} \\ J_{31} & J_{32} & J_{33} \end{pmatrix} \tag{15}$$

where

$$\begin{aligned} J_{11} &= (1-2x)\{(\psi_{P1}+\psi_{P2})y + [\beta F + \beta\theta_2 pD(1+i)]z \\ &+ (\alpha-\beta)[F+\theta_2 pD(1+i)]yz + C_{P2} - C_{P1} \\ &- \theta_1 pD(1+i)\} \end{aligned} \tag{16}$$

$$\begin{aligned} J_{12} &= x(1-x)\{\psi_{P1}+\psi_{P2} \\ &+ (\alpha-\beta)[F+\theta_2 pD(1+i)]z\} \end{aligned} \tag{17}$$

$$\begin{aligned} J_{13} &= x(1-x)\{\beta F + \beta\theta_2 pD(1+i) \\ &+ (\alpha-\beta)[F+\theta_2 pD(1+i)]y\} \end{aligned} \tag{18}$$

$$J_{21} = y(1-y)(\alpha-\beta)(1-\theta_2)pD(1+i)z \tag{19}$$

$$\begin{aligned} J_{22} &= (1-2y)\{[R_L - (\alpha-\beta)(1-\theta_2)pD(1+i)]z \\ &+ (\alpha-\beta)(1-\theta_2)pD(1+i)xz - C_L/D\} \end{aligned} \tag{20}$$

$$\begin{aligned} J_{23} &= y(1-y)[R_L - (\alpha-\beta)(1-\theta_2)pD(1+i) \\ &+ (\alpha-\beta)(1-\theta_2)pD(1+i)x] \end{aligned} \tag{21}$$

$$J_{31} = -z(1-z)\{-\beta(F+\psi_{W2}) - [(\alpha-\beta)(F+\psi_{W2})]y\} \tag{22}$$

$$\begin{aligned} J_{32} &= z(1-z)[\psi_{R1}+\psi_{R2}+(\alpha-\beta)(F+\psi_{W2}) - R_L \\ &- (\alpha-\beta)(F+\psi_{W2})x] \end{aligned} \tag{23}$$

$$\begin{aligned} J_{33} &= (1-2z)\{-\beta(F+\psi_{W2})x + [\psi_{R1}+\psi_{R2} \\ &+ (\alpha-\beta)(F+\psi_{W2}) - R_L]y - [(\alpha-\beta)(F+\psi_{W2})]xy \\ &+ C_{R2} + \beta(F+\psi_{W2}) - C_{R1} - M_R\} \end{aligned} \tag{24}$$

Substitute the eight pure strategy Nash equilibrium points in the Jacobian matrix (15), the corresponding eigenvalues can be obtained, and thus the asymptotic stability of the eight equilibrium points can be judged.

(1) For the equilibrium point $E_1(0,0,0)$, the eigenvalues of the Jacobian matrix are as follows:

$$\begin{cases} \lambda_1^1 = C_{P2} - C_{P1} - \theta_1 pD(1+i) \\ \lambda_2^1 = -C_L/D \\ \lambda_3^1 = C_{R2} + \beta(F+\psi_{W2}) - C_{R1} - M_R \end{cases} \tag{25}$$

It is noted that $\lambda_1^1 = C_{P2} - C_{P1} - \theta_1 pD(1+i) < 0$,$\lambda_2^1 = -C_L/D < 0$, thus, according to the judgment criterion, when $C_{R2} + \beta(F+\psi_{W2}) - C_{R1} - M_R < 0$,$E_1(0,0,0)$ will be an ESS, in this case, the P2P lending platforms will adopt the "negative positive" strategy, the investors will adopt the "not participating in co-governance" strategy, and the financial regulators will adopt the "traditional regulation" strategy, which is the worst state.

(2) For the equilibrium point $E_2(1,0,0)$, the eigenvalues of the Jacobian matrix are as follows:

$$\begin{cases} \lambda_1^2 = -[C_{P2} - C_{P1} - \theta_1 pD(1+i)] \\ \lambda_2^2 = -C_L/D \\ \lambda_3^2 = C_{R2} - C_{R1} - M_R \end{cases} \tag{26}$$

According to the assumptions, $\lambda_1^2 = -[C_{P2} - C_{P1} - \theta_1 pD(1+i)] > 0$, thus, the judgment criterion is not satisfied, and $E_2(1,0,0)$ cannot be an ESS.

(3) For the equilibrium point $E_3(0,1,0)$, the eigenvalues of the Jacobian matrix are as follows:

$$\begin{cases} \lambda_1^3 = \psi_{P1} + \psi_{P2} + C_{P2} - C_{P1} - \theta_1 pD(1+i) \\ \lambda_2^3 = C_L/D \\ \lambda_3^3 = \psi_{R1} + \psi_{R2} + \alpha(F+\psi_{W2}) - R_L + C_{R2} - C_{R1} - M_R \end{cases} \tag{27}$$

where C_L/D refers to the cost of investors when adopting the "participating in co-governance" strategy and cannot be negative, thus $\lambda_2^3 > 0$ does not meet the judgment criterion, and $E_3(0,1,0)$ cannot be an ESS.

(4) For the equilibrium point $E_4(0,0,1)$, the eigenvalues of the Jacobian matrix are as follows:

$$\begin{cases} \lambda_1^4 = \beta F + \beta\theta_2 pD(1+i) + C_{P2} - C_{P1} - \theta_1 pD(1+i) \\ \lambda_2^4 = R_L - \frac{C_L}{D} - (\alpha-\beta)(1-\theta_2)pD(1+i) \\ \lambda_3^4 = -[C_{R2} + \beta(F+\psi_{W2}) - C_{R1} - M_R] \end{cases} \tag{28}$$

when $\beta F + \beta\theta_2 pD(1+i) + C_{P2} - C_{P1} - \theta_1 pD(1+i) < 0$,$R_L - C_L/D - (\alpha-\beta)(1-\theta_2)pD(1+i) < 0$ and $C_{R2} + \beta(F+\psi_{W2}) - C_{R1} - M_R > 0$, $E_4(0,0,1)$ will be an ESS.

(5) For the equilibrium point $E_5(0,1,1)$, the eigenvalues of the Jacobian matrix are as follows:

$$\begin{cases} \lambda_1^5 = \psi_{P1} + \psi_{P2} + \alpha F + \alpha\theta_2 pD(1+i) + C_{P2} - C_{P1} - \theta_1 pD(1+i) \\ \lambda_2^5 = -\left[R_L - \frac{C_L}{D} - (\alpha-\beta)(1-\theta_2)pD(1+i)\right] \\ \lambda_3^5 = -[\psi_{R1} + \psi_{R2} + \alpha(F+\psi_{W2}) - R_L + C_{R2} - C_{R1} - M_R] \end{cases} \tag{29}$$

when $\psi_{P1} + \psi_{P2} + \alpha F + \alpha\theta_2 pD(1+i) + C_{P2} - C_{P1} - \theta_1 pD(1+i) < 0$, $R_L - C_L/D - (\alpha-\beta)(1-\theta_2)pD(1+i) > 0$ and $\psi_{R1} + \psi_{R2} + \alpha(F+\psi_{W2}) - R_L + C_{R2} - C_{R1} - M_R > 0$,$E_5(0,1,1)$ will be an ESS.

(6) For the equilibrium point $E_6(1,1,0)$, the eigenvalues of the Jacobian matrix are as follows:

TABLE 4 Stability conditions of the equilibrium points.

Equilibrium point	Conditions of ESS
$E_1(0,0,0)$	$C_{R2}+\beta(F+\psi_{W2})-C_{R1}-M_R<0$
$E_4(0,0,1)$	$\beta F+\beta\theta_2 pD(1+i)+C_{P2}-C_{P1}-\theta_1 pD(1+i)<0, R_L-C_L-(\alpha-\beta)(1-\theta_2)pD(1+i)<0,$ $C_{R2}+\beta(F+\psi_{W2})-C_{R1}-M_R>0$
$E_5(0,1,1)$	$\psi_{P1}+\psi_{P2}+\alpha F+\alpha\theta_2 pD(1+i)+C_{P2}-C_{P1}-\theta_1 pD(1+i)<0,$ $R_L-C_L-(\alpha-\beta)(1-\theta_2)pD(1+i)>0,$ $\psi_{R1}+\psi_{R2}+\alpha(F+\psi_{W2})-R_L+C_{R2}-C_{R1}-M_R>0$
$E_8(1,1,1)$	$\psi_{P1}+\psi_{P2}+\alpha F+\alpha\theta_2 pD(1+i)+C_{P2}-C_{P1}-\theta_1 pD(1+i)>0,$ $C_L-R_L<0,$ $\psi_{R1}+\psi_{R2}-R_L+C_{R2}-C_{R1}-M_R>0$

$$\begin{cases}\lambda_1^6=-[\psi_{P1}+\psi_{P2}+C_{P2}-C_{P1}-\theta_1 pD(1+i)]\\ \lambda_2^6=C_L/D\\ \lambda_3^6=\psi_{R1}+\psi_{R2}-R_L+C_{R2}-C_{R1}-M_R\end{cases}\tag{30}$$

where $C_L$refers to the cost of investors when adopting the "participating in co-governance" strategy and cannot be negative, thus $\lambda_2^6>0$ does not satisfy the judgment criterion, and $E_6(1,1,0)$ cannot be an ESS.

(7) For the equilibrium point $E_7(1,0,1)$, the eigenvalues of the Jacobian matrix are as follows:

$$\begin{cases}\lambda_1^7=-[\beta F+\beta\theta_2 pD(1+i)+C_{P2}-C_{P1}-\theta_1 pD(1+i)]\\ \lambda_2^7=R_L-C_L/D\\ \lambda_3^7=-(C_{R2}-C_{R1}-M_R)\end{cases}\tag{31}$$

Since $\lambda_3^7=-(C_{R2}-C_{R1}-M_R)>0$ does not satisfy the judgment criterion, thus, $E_7(1,0,1)$ cannot be an ESS.

(8) For the equilibrium point $E_8(1,1,1)$, the eigenvalues of the Jacobian matrix are as follows:

$$\begin{cases}\lambda_1^8=-[\psi_{P1}+\psi_{P2}+\alpha F+\alpha\theta_2 pD(1+i)\\ \quad +C_{P2}-C_{P1}-\theta_1 pD(1+i)]\\ \lambda_2^8=C_L/D-R_L\\ \lambda_3^8=-[\psi_{R1}+\psi_{R2}-R_L+C_{R2}-C_{R1}-M_R]\end{cases}\tag{32}$$

When $\psi_{P1}+\psi_{P2}+\alpha F+\alpha\theta_2 pD(1+i)+C_{P2}-C_{P1}-\theta_1 pD(1+i)>0$, $C_L/D-R_L<0$ and $\psi_{R1}+\psi_{R2}-R_L+C_{R2}-C_{R1}-M_R>0$, $E_8(1,1,1)$ will be an ESS.

Overall, the equilibrium points $E_2(1,0,0)$, $E_3(0,1,0)$, $E_6(1,1,0)$, and $E_7(1,0,1)$ cannot be ESSs under any conditions, while $E_1(0,0,0)$,$E_4(0,0,1)$,$E_5(0,1,1)$, and $E_8(1,1,1)$ can be ESSs under certain conditions. Based on the above analysis, the stability conditions of the four equilibrium points, which may be an ESS, are listed in Table 4.

Comparing the four possible ESSs, from the perspective of consumer protection and the sustainable development of the FinTech industry, $E_8(1,1,1)$ is the most appropriate ESS, i.e., P2P lending platforms adopt the "positive disposal" strategy, investors adopt the "participating in co-governance" strategy, and financial regulators adopt the "co-governance regulation" strategy. The reasons are as follows: First, P2P lending platforms will be eager to adopt a "positive disposal" strategy, as this will improve their reputation and establish a solid foundation for the future development of the FinTech industry. Second, the fundamental purpose of the co-governance regulation is to protect the consumer, only when the investors adopt a "participating in co-governance" strategy, do they have the possibility to take their money back quickly and smoothly. Third, the current benign exit of P2P lending platforms in China is inseparable from the regulation of the financial regulators, if the investors can cooperate with the financial regulators to form a social co-governance pattern with high regulatory efficiency, then the benign exit of P2P lending platforms will be accelerated, and the rights and interests of the consumer will be ensured.

To guarantee that $E_8(1,1,1)$ is the only evolutionary stable strategy of the tripartite game, the stability constraints are listed as follows:

(1) $\lambda_1^8<0$, $\lambda_2^8<0$ and $\lambda_3^8<0$. Correspondingly, $\psi_{P1}+\psi_{P2}+\alpha F+\alpha\theta_2 pD(1+i)+C_{P2}-C_{P1}-\theta_1 pD(1+i)>0$ means that for P2P lending platforms the payoffs of adopting a "positive disposal" strategy with reputation benefits should be higher than that of adopting a "negative disposal" strategy after suffering a reputation loss and being punished by the financial regulators. $C_L/D-R_L<0$ means that for investors the rewards of adopting the "participating in co-governance" strategy should be higher than the costs. $\psi_{R1}+\psi_{R2}-R_L+C_{R2}-C_{R1}-M_R>0$ means that for financial regulators, the payoff difference between co-governance regulation and traditional regulation should be higher than the sum of the initiative cost to set up the social co-governance pattern and the rewards paid to the investors. The three conditions need to be satisfied simultaneously.

(2) $\lambda_3^1 > 0$. Correspondingly, $C_{R2}+\beta(F+\psi_{W2})-C_{R1}-M_R > 0$ means that for financial regulators the cost of adopting "traditional regulation" is higher than that of adopting "co-governance regulation."

(3) $\lambda_1^4 > 0$, or $\lambda_2^4 > 0$. Correspondingly, $\beta F + \beta\theta_2 pD(1 + i) + C_{P2} - C_{P1} - \theta_1 pD(1 + i) > 0$ means that for P2P lending platforms the cost of adopting aa "negative disposal" strategy after being punished by the financial regulators should of higher than that of adopting a "positive disposal" strategy. $R_L - C_L/D - (\alpha - \beta)(1-\theta_2)pD(1+i) > 0$ means that for investors the payoffs of "participating in co-governance" should be higher than that of "not participating in co-governance." However, since $\lambda_3^4 = -\lambda_3^1$, thus, when condition (2) holds, λ_3^4 cannot be positive.

(4) $\lambda_1^5 > 0$, or $\lambda_2^5 > 0$, or $\lambda_3^5 > 0$. From the above analysis, $\lambda_1^5 = -\lambda_1^8$, thus, when condition (1) is satisfied, then condition (2) holds.

To guarantee that $E_8(1, 1, 1)$ is the only evolutionary stable strategy of the tripartite game, the above four conditions must be held at the same time.

Numerical simulation

Through the theoretical analysis illustrated above, four ESSs have been identified, which can be obtained when corresponding conditions are satisfied. Moreover, to intuitively observe the evolutionary trajectories of the stakeholders and their sensitivity to parameters, it is necessary to simulate their strategies. In this study, we implemented this simulation by using MATLAB.

Evolutionary stable strategies

Scenario 1

In this scenario, when $C_{R2} + \beta(F + \psi_{W2}) - C_{R1} - M_R < 0$, $E_1(0, 0, 0)$ will be the ESS. In order to meet these conditions, suppose that $C_{P1} = 3$, $C_{P2} = 1$, $\theta_1 = 0.9$, $\theta_2 = 0.2$, $F = 1$, $\psi_{P1} = 0.5$, $\psi_{P2} = 1$, $R_L = 0.5$, $C_L = 0.5$, $D = 0.5$, $P = 10$, $i = 0.1$, $C_{R1} = 5.5$, $C_{R2} = 3.5$, $\alpha = 0.8$, $\beta = 0.5$, $\psi_{w1} = 4$, $\psi_{w2} = 6$, $\psi_{R1} = 2$, $\psi_{R2} = 4$ and $M_R = 5$. Then, as shown in Figure 3A, regardless of the initial proportion of the three participants, the definitive evolutionary result is $E_1(0, 0, 0)$. The main reason for this result is that the payoff of P2P lending platforms' "negative disposal" strategy is higher than that of "positive disposal" due to the limited penalty of financial regulators to the P2P lending platforms and the lack of reputation effect, while the payoff of investors participating in co-governance is lower than their basic earning, and the reputation benefits of the financial regulators are not enough to offset the difference between the fine and co-governance regulation cost. Thus, the ESS of P2P lending platforms, investors, and financial regulators is "negative disposal," "not participating in co-governance," and "traditional regulation," respectively.

Scenario 2

In this scenario, when $\beta F + \beta\theta_2 pD(1 + i) + C_{P2} - C_{P1} - \theta_1 pD(1 + i) < 0$, $R_L - C_L/D - (\alpha - \beta)(1 - \theta_2)pD(1 + i) < 0$ and $C_{R2} + \beta(F + \psi_{W2}) - C_{R1} - M_R > 0$, the equilibrium point $E_4(0, 0, 1)$ will be the ESS. To meet these conditions, suppose that $C_{P1} = 3$,$C_{P2} = 1$,$\theta_1 = 0.9$, $\theta_2 = 0.2$, $F = 4$, $\psi_{P1} = 0.5$, $\psi_{P2} = 1$, $R_L = 1$, $C_L = 0.5$, $D = 0.5$, $P = 10$, $i = 0.1$, $C_{R1} = 5.5$, $C_{R2} = 3.5$, $\alpha = 0.8$, $\beta = 0.5$, $\psi_{w1} = 4$, $\psi_{w2} = 6$, $\psi_{R1} = 3$, $\psi_{R2} = 5$ and $M_R = 2$. Then, Figure 3B shows that the proportion of P2P lending platforms adopting the "positive disposal" strategy, and the proportion of investors adopting the "not participating in co-governance" strategy constantly decreases with the iteration of the evolution, and eventually converges to 0, whereas the proportion of financial regulators adopting "co-governance regulation" continually increases, and eventually converges to 1. Clearly, the reasons why P2P lending platforms adopt such a strategy is similar to scenario 1, and since the rewards of financial regulators to investors are not attractive to offset the costs, then the investors will still adopt the "not participating in co-governance" strategy. However, the payoff of financial regulators' "co-governance regulation" is higher than that of the "traditional regulation" strategy; thus, in this scenario, financial regulators are more willing to adopt the "co-governance regulation" strategy.

Scenario 3

In this scenario, when $\psi_{P1}+\psi_{P2}+\alpha F+\alpha\theta_2 pD(1+i)+C_{P2}-C_{P1}-\theta_1 pD(1+i) < 0$, $R_L-C_L/D-(\alpha-\beta)(1-\theta_2)pD(1+i) > 0$ and $\psi_{R1} + \psi_{R2} + \alpha(F + \psi_{W2}) - R_L + C_{R2} - C_{R1} - M_R > 0$, the equilibrium point $E_5(0, 1, 1)$ is the ESS. To meet these conditions, suppose that $C_{P1} = 3$, $C_{P2} = 1$, $\theta_1 = 0.9$, $\theta_2 = 0.2$, $F = 5$, $\psi_{P1} = 0.5$, $\psi_{P2} = 1$, $R_L = 3$, $C_L = 0.5$, $D = 0.5$, $P = 10$, $i = 0.1$, $C_{R1} = 5.5$, $C_{R2} = 3.5$, $\alpha = 0.8$, $\beta = 0.5$, $\psi_{w1} = 4$, $\psi_{w2} = 6$, $\psi_{R1} = 3$, $\psi_{R2} = 5$ and $M_R = 2$. Then, Figure 3C shows that the ESS will eventually stabilize at the equilibrium point $E_5(0, 1, 1)$. This phenomenon demonstrates that the financial regulators will adopt the "co-governance regulation" strategy due to the increase in their payoffs, and the investors will also adopt the "co-governance" strategy due to the increase in rewards. However, since there does not exist solid reputation mechanism for the P2P lending platforms and the penalty is not properly formulated, thus, after comparing the payoffs of the "positive disposal" and "negative disposal" strategy, the P2P lending platforms will eventually adopt the "negative disposal" strategy.

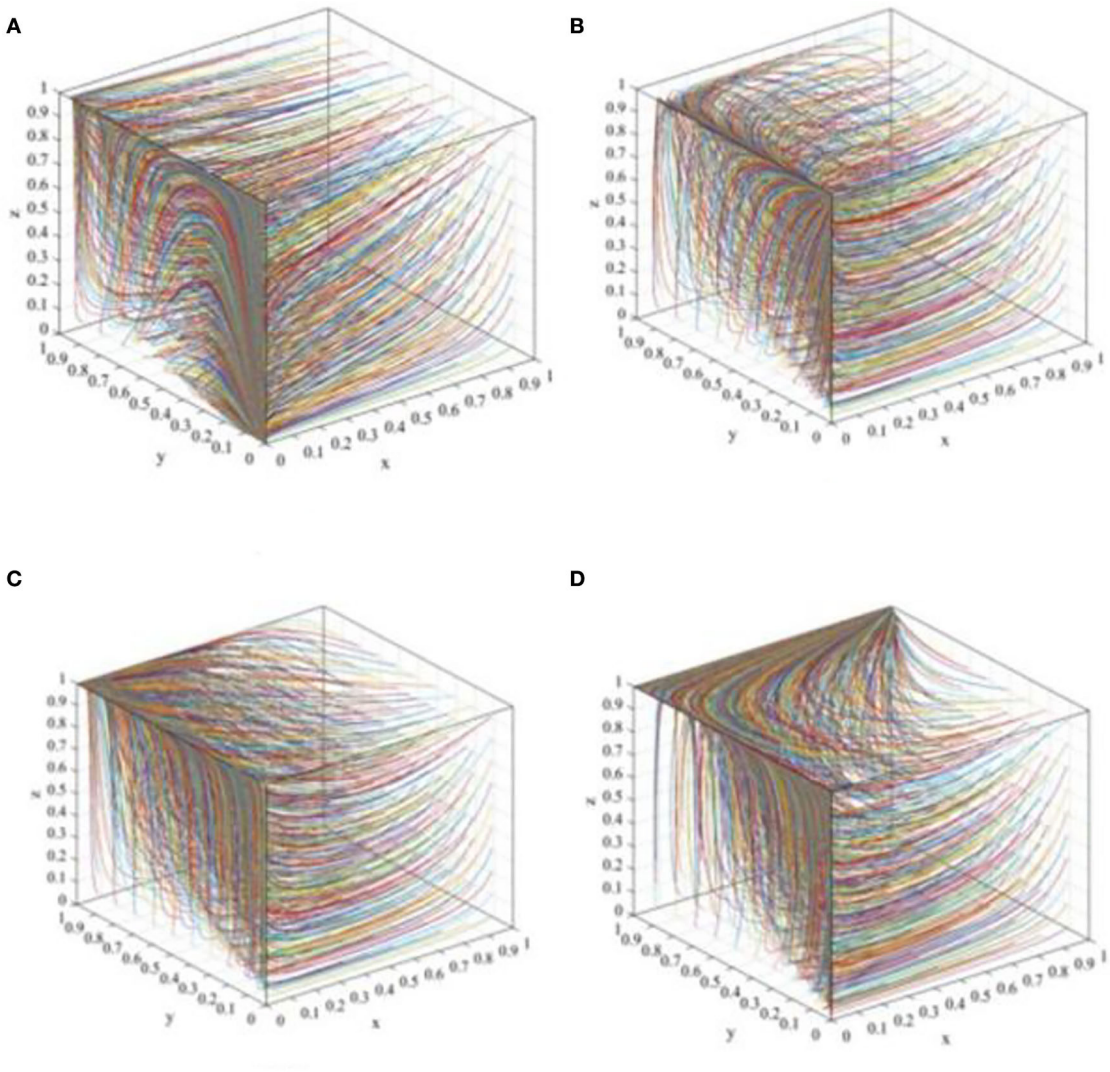


FIGURE 3
The evolutionary path diagram of the three participants in different scenarios. **(A)** The evolutionary path at the equilibrium point $E_1(0, 0, 0)$; **(B)** The evolutionary path at the equilibrium point $E_4(0, 0, 1)$; **(C)** The evolutionary path at the equilibrium point $E_5(0, 1, 1)$; and **(D)** The evolutionary path at the equilibrium point $E_8(1, 1, 1)$.

Scenario 4

In this scenario, when $\psi_{P1} + \psi_{P2} + \alpha F + \alpha\theta_2 pD(1 + i) + C_{P2} - C_{P1} - \theta_1 pD(1 + i) > 0, C_L/D - R_L < 0$ and $\psi_{R1} + \psi_{R2} - R_L + C_{R2} - C_{R1} - M_R > 0$, the equilibrium point $E_8(1, 1, 1)$ is the ESS. To meet these conditions, suppose that $C_{P1} = 3$, $C_{P2} = 1$, $\theta_1 = 0.9$, $\theta_2 = 0.2$, $F = 5$, $\psi_{P1} = 2$, $\psi_{P2} = 3$, $R_L = 3$, $C_L = 0.5$, $D = 0.5$, $P = 10$, $i = 0.1$, $C_{R1} = 5.5$, $C_{R2} = 3.5$, $\alpha = 0.8$, $\beta = 0.5$, $\psi_{w1} = 4$, $\psi_{w2} = 6$, $\psi_{R1} = 3$, $\psi_{R2} = 5$ and $M_R = 2$. Then, as shown in Figure 3D, the proportion of P2P lending platforms adopting the "positive disposal" strategy, the proportion of investors adopting the "not participating in co-governance" strategy, and the proportion of financial regulators adopting "co-governance regulation" continually increase and eventually converges to 1. In this case, the financial regulators are more willing to carry out co-governance regulation due to the high reputation benefits from the investors and the decrease in the initiative cost to set up the social co-governance pattern. Accordingly, the investors will adopt the "co-governance" strategy due to the high rewards. Consequently, if the P2P lending platforms adopt the "negative disposal" strategy, then they need not only to pay corresponding fines but also suffer a great loss of their reputation, which is fatal to their future development; thus, the P2P lending platforms will eventually adopt the "positive disposal" strategy.

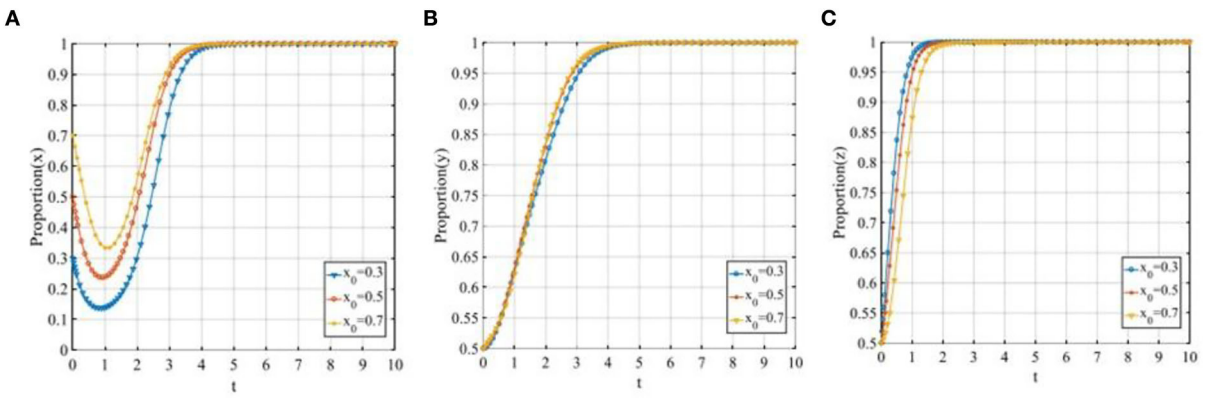


FIGURE 4
The evolutionary trajectories of three participants when $x_0 = 0.3,\ 0.5,\ 0.7$. **(A)** The sensitivity of P2P lending platforms; **(B)** The sensitivity of investors; and **(C)** The sensitivity of financial regulators.

Sensitivity analysis

From the above analysis, the equilibrium point $E_8(1, 1, 1)$ is the ideal ESS of the tripartite evolutionary model, in order to more intuitively illustrate the influence of the changes in parameters on the stable state of the tripartite evolutionary game, this subsection will use numerical simulations to explore the sensitivity of the three participants to some key parameters, i.e., the initial proportion of three participants, the cost of P2P lending platforms when adopting "positive disposal" strategy (C_{P1}), the penalty of financial regulators to P2P lending platforms (F), the reputation effects of P2P lending platforms (ψ_{P1},ψ_{P2}), the rewards of financial regulators to investors (R_L), the cost of investors (C_L), the centrality degree (D), the investment volume of investors (P), the reputation benefits of financial regulators (ψ_{R1}), and the initiative cost of financial regulators to set up the co-governance regulation pattern (M_R). It should be noted that when we analyze the sensitivity of one of these parameters, the remaining parameters' values remain unchanged. The initial settings of the parameters are the same as in scenario 4.

The initial proportion of three participants

To further illustrate the influence of the initial proportion of three participants on the tripartite evolutionary game, we let x_0 equals 0.3, 0.5, and 0.7, y_0 equals 0.3, 0.5, and 0.7, z_0 also equals 0.3, 0.5, and 0.7 in turn. It is noteworthy that the initial states of the remaining stakeholders are controlled at 0.5 when we probe the sensitivity of each stakeholder to the initial state.

Set $x_0 = 0.3,\ 0.5,\ 0.7$, the evolutionary trajectories of the three participants are shown in Figure 4. The change in the initial proportion of P2P lending platforms has a significant impact on P2P lending platforms, the higher the initial proportion is, the faster the P2P lending platforms will converge to 1. While, as shown in Figures 4B,C, the change in the initial proportion of P2P lending platforms has a limited effect on investors and financial regulators.

Set $y_0 = 0.3,\ 0.5,\ 0.7$, the evolution of the three participants is shown in Figure 5. The change in the initial proportion of investors has a significant impact on P2P lending platforms and investors while having a limited effect on financial regulators. With the increase in the initial proportion of the investors adopting the "participating in co-governance" strategy, the convergence rates of the P2P lending platforms and investors will also increase, the higher the initial proportion is, the faster the P2P lending platforms and investors will converge to 1.

Set $z_0 = 0.3,\ 0.5,\ 0.7$, the evolutionary trajectories of the three participants are shown in Figure 6. The change in the initial proportion of financial regulators has a significant impact on P2P lending platforms, investors, and financial regulators, respectively. With the increase in the initial proportion of the financial regulators adopting the "co-governance regulation" strategy, the convergence rates of the three participants will all increase, the higher the initial proportion is, the faster the evolutionary trajectories will finally stabilize at the ideal state $E_8(1, 1, 1)$.

From the above analysis, we can see that the initial proportion of the three participants will impact the speed to reach the final ESS but would not change the final stable state. Therefore, we set $x_0 = y_0 = z_0 = 0.5$ to further illustrate the sensitivity of other parameters in the following sections.

The cost of P2P lending platforms' "positive disposal" strategy

To explore the sensitivity of three participants to the cost of P2P lending platforms when adopting a "positive disposal"

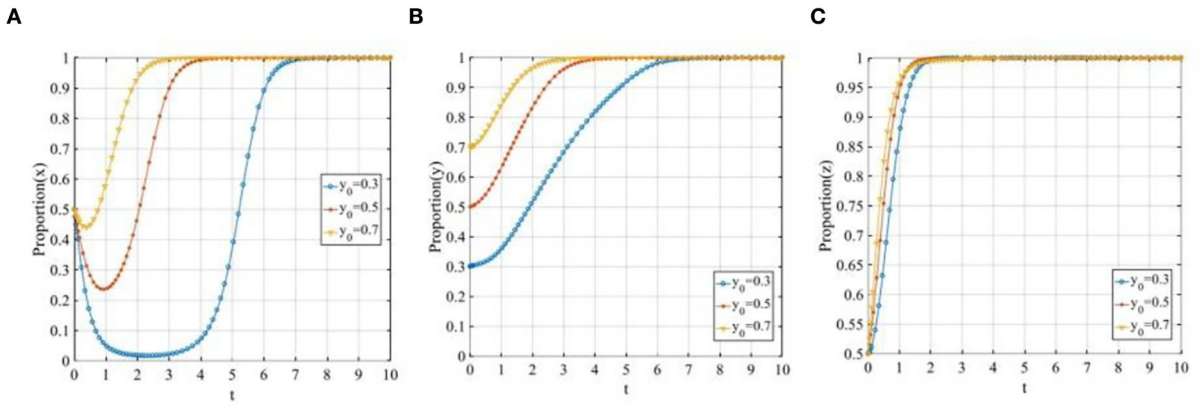


FIGURE 5
The evolutionary trajectories of three participants when $y_0 = 0.3,\ 0.5,\ 0.7$. **(A)** The sensitivity of P2P lending platforms; **(B)** The sensitivity of investors; and **(C)** The sensitivity of financial regulators.

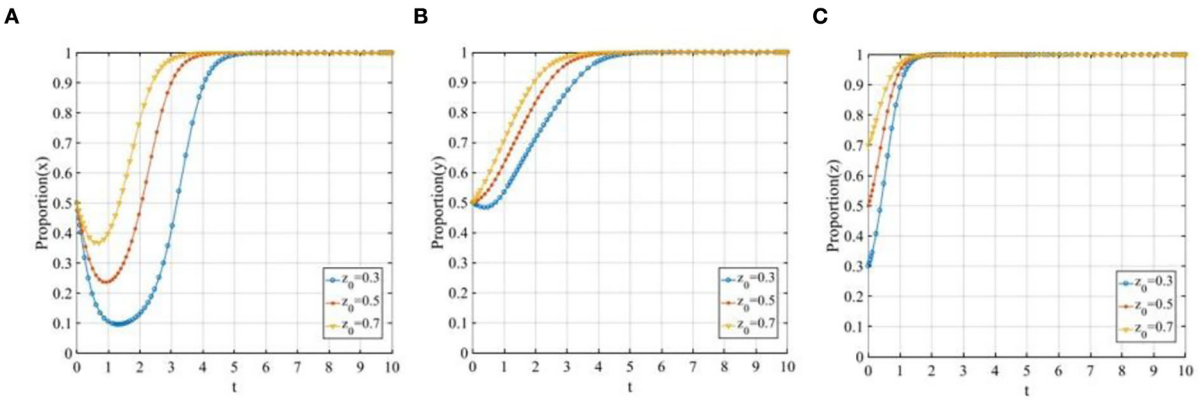


FIGURE 6
The evolutionary trajectories of three participants when $z_0 = 0.3,\ 0.5,\ 0.7$. **(A)** The sensitivity of P2P lending platforms; **(B)** The sensitivity of investors; and **(C)** The sensitivity of financial regulators.

strategy, let C_{P1} equals 1, 2, 3, 4, 5, and 6, respectively. The evolutionary trajectories of the three participants are shown in Figure 7. As seen from Figure 7A, P2P lending platforms are more willing to adopt the "positive disposal" strategy when the costs are relatively low. However, with the increase in costs, the evolutionary speed of the P2P lending platforms will be slower, meanwhile, once the costs exceed the threshold, P2P lending platforms will change their strategy from 1 to 0. Figure 7B shows that with the increase in costs, although investors will eventually adopt the "participating in co-governance" strategy, the time to reach the stable status will be continually prolonged. Figure 7C shows that with the increase in costs, the evolutionary speed of financial regulators will be accelerated.

The penalty from financial regulators for P2P lending platforms

To explore the sensitivity of three participants to the penalty of financial regulators to P2P lending platforms, let F equals to 1, 2, 3, 4, 5, and 6. The evolutionary trajectories of the three participants are shown in Figure 8. As seen from Figure 8A, it is obvious that if the penalty is not properly formulated, the P2P lending platforms will eventually adopt the "negative disposal" strategy, when the penalty increases gradually, the P2P lending platforms will tend to adopt the "positive disposal" strategy, and with the increase in the penalty, the evolutionary speed will be improved sharply. Figure 8B shows that with the increase in the penalty, the evolutionary speed of the investors adopting "participating in co-governance" will be accelerated. Figure 8C

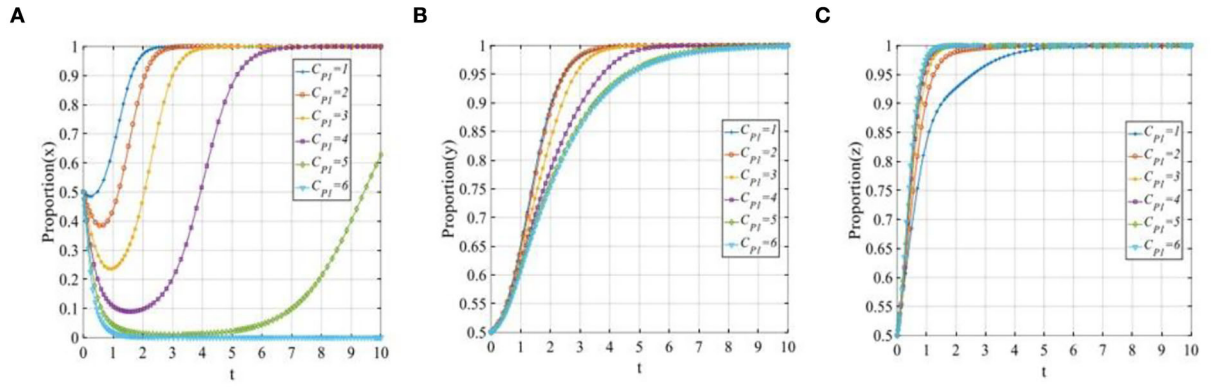


FIGURE 7
The sensitivity analysis of the cost of P2P lending platforms' "positive disposal" strategy when $C_{P1} = 1, 2, 3, 4, 5, 6$. **(A)** The sensitivity of P2P lending platforms; **(B)** The sensitivity of investors; and **(C)** The sensitivity of financial regulators.

shows that the evolutionary speed of the financial regulators will also be accelerated, but the effect is not significant compared with the other two participants.

The reputation effects of P2P lending platforms

To explore the sensitivity of three participants to the reputation benefits of P2P lending platforms, let ψ_{P1} equals 0.5, 1, 1.5, 2, 2.5, and 3. The evolutionary trajectories of the three participants are shown in Figure 9. With the increase of the reputation benefits of the P2P lending platforms adopting a "positive disposal" strategy, the convergence rates of the P2P lending platforms and investors will also increase, the higher the reputation benefit is, the faster the P2P lending platforms and investors will converge to 1. However, in Figure 9C, the nearly overlapping curves demonstrate that the penalty has little effect on the financial regulators.

To explore the sensitivity of three participants to the reputation loss of P2P lending platforms, let ψ_{P2} equals 0, 2, 2.5, 3, 3.5, and 4. The evolutionary trajectories of the three participants are shown in Figure 10. As seen from Figure 10A, P2P lending platforms tend to adopt the "positive disposal" strategy except for a sudden change in reputation loss ($\psi_{P2} = 0$). This means that when the reputation losses exceed the threshold, the P2P lending platforms will adopt the "positive disposal" strategy, and the evolutionary speed will be improved with the increase in the reputation losses. However, if the reputation losses of P2P lending platforms adopting a "negative disposal" strategy are lower than the threshold, then the P2P lending platforms will change their strategy from 1 to 0. Figure 10B shows that with the continual increase in reputation losses, the time of the investors adopting the "participating in co-governance" strategy will be continually shortened. In Figure 10C, similarly to the case of reputation benefits, reputation losses have little significant effect on the financial regulators.

The rewards of financial regulators to investors

To explore the sensitivity of three participants to the rewards of financial regulators to investors, let R_L equals 0, 1, 2, 3, 4, and 5. The evolutionary trajectories of the three participants are shown in Figure 11. The change in the rewards of financial regulators to investors has a significant impact on P2P lending platforms, investors, and financial regulators, respectively. Figure 11A shows that with the increase in the rewards, the P2P lending platforms will change their strategy from 0 to 1. Figure 11B shows that if the rewards are small, the investor is not willing to adopt the "participating in co-governance" strategy, once the rewards exceed the threshold, the investors will tend to adopt the "participating in co-governance" strategy, and the evolutionary time will be shorted with the increase in the rewards. Interestingly, Figure 11C shows that the sensitivity of the financial regulators to the rewards is opposite to those of the P2P lending platforms and the investors, once the rewards exceed the threshold, the financial regulators will tend to adopt the "traditional regulation" strategy.

The cost to investors

To explore the sensitivity of three participants to the cost of investors when adopting the "participating in co-governance" strategy, let C_L equals 0, 0.5, 1, 1.5, 2, and 4. The evolutionary trajectories of the three participants are shown in Figure 12. With the increase in the cost of the investors adopting the "participating in co-governance" strategy, the P2P lending platforms and investors will change their strategy from 1 to 0, and the higher the costs of investors are, the faster the P2P

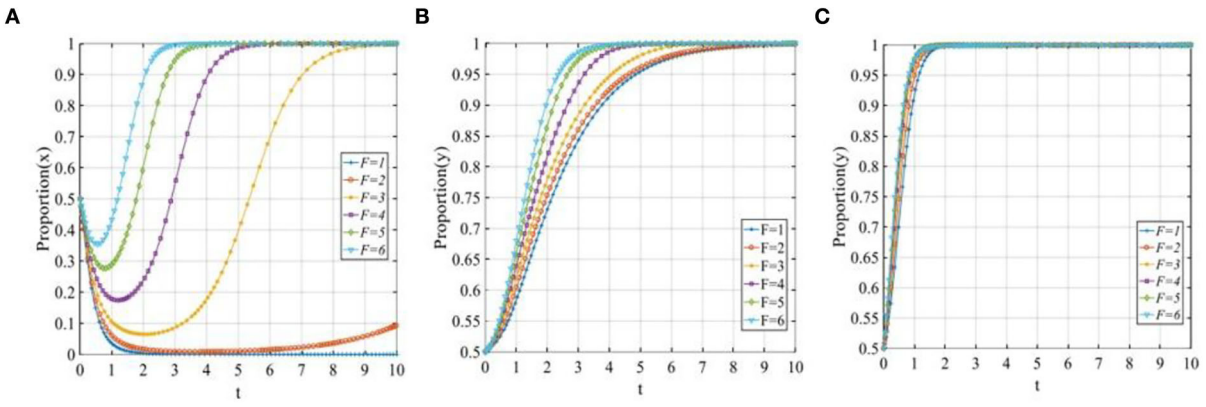


FIGURE 8
The sensitivity analysis of the penalty of financial regulators to P2P lending platforms when $F = 1, 2, 3, 4, 5, 6$. **(A)** The sensitivity of P2P lending platforms; **(B)** The sensitivity of investors; and **(C)** The sensitivity of financial regulators.

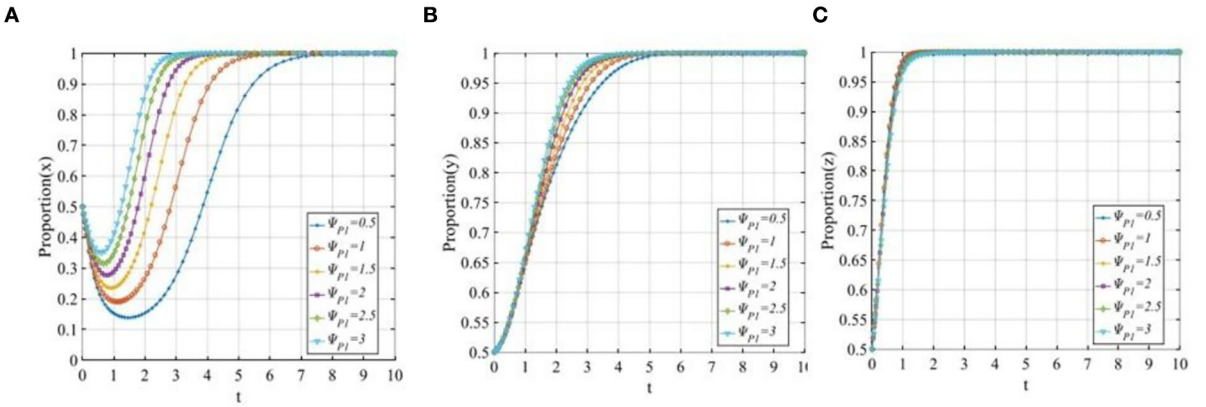


FIGURE 9
The sensitivity analysis of the reputation benefits of P2P lending platforms when $\psi_{P1} = 0.5, 1, 1.5, 2, 2.5, 3$. **(A)** The sensitivity of P2P lending platforms; **(B)** The sensitivity of investors; and **(C)** The sensitivity of financial regulators.

lending platforms and investors will converge to 0. Additionally, Figure 12C shows that with the increase in the costs of investors, although financial regulators will eventually choose the "co-governance regulation" strategy, the time to reach a stable status will be continually prolonged.

The centrality degree of investors

Network centrality plays a critical role in shaping lenders' investment behavior. According to the prior study (Chen et al., 2022), lenders who are in the center of a network not only invest by larger amounts but also more swiftly than their peers, reflecting the experience and information advantage arising from their position in the network. To explore the sensitivity of three participants toward the network centrality of investors, we use the centrality degree D (ranges from 0 to 1) (Freeman, 1978) to measure the network centrality of investors (Chen et al., 2022). We present sensitivity analysis for D equals 0.1, 0.2, 0.3, 0.4, 0.5, and 0.6, respectively. The evolutionary trajectories of the three participants are shown in Figure 13. From Figure 13A, we can see that with the increase in the centrality degree of investors (invest by higher fractions and more swiftly), the P2P lending platforms will change their strategy from 0 to 1. Similarly, Figure 13B shows that if the centrality degree of investors is low, the investors are not willing to adopt the "participating in co-governance" strategy. However, when the centrality degree of investors exceeds the threshold ($D = 0.2$), the investors would tend to adopt the "participating in co-governance" strategy, and the evolutionary time would be reduced with the increase in the centrality degree of investors.

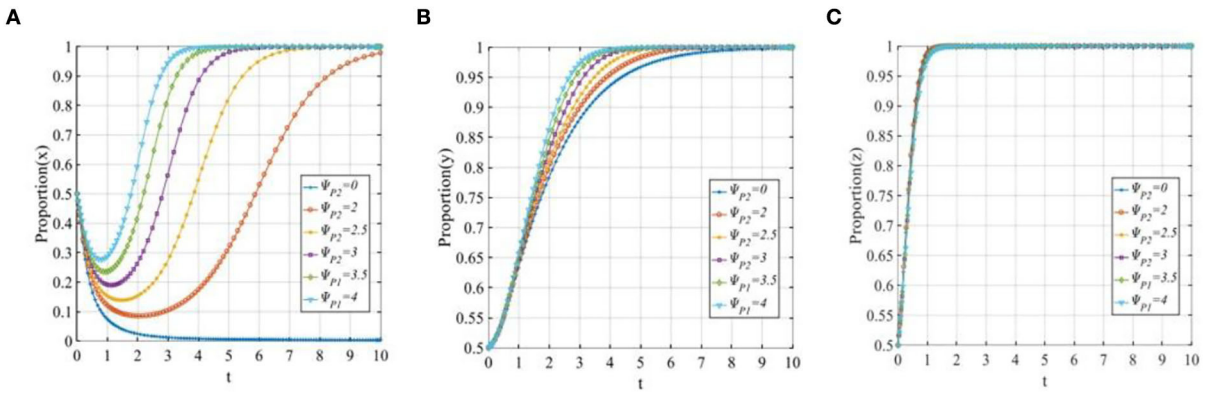


FIGURE 10
The sensitivity analysis of the reputation losses of P2P lending platforms when $\psi_{P2} = 0,\ 2,\ 2.5,\ 3,\ 3.5,\ 4$. **(A)** The sensitivity of P2P lending platforms; **(B)** The sensitivity of investors; and **(C)** The sensitivity of financial regulators.

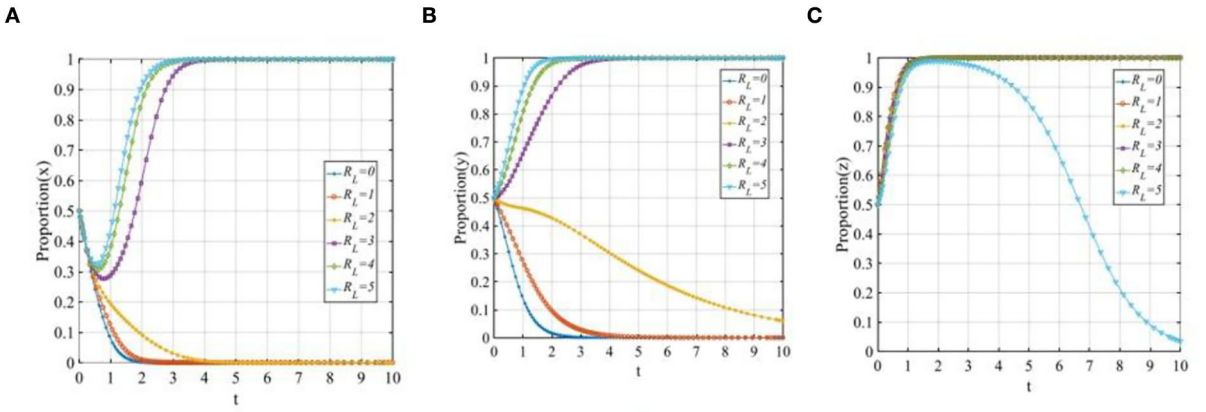


FIGURE 11
The sensitivity analysis of the rewards of financial regulators to investors when $R_L = 0,\ 1,\ 2,\ 3,\ 4,\ 5$. **(A)** The sensitivity of P2P lending platforms; **(B)** The sensitivity of investors; and **(C)** The sensitivity of financial regulators.

Figure 13C shows that the financial regulators' evolutionary speed to adopt the "participating in co-governance" strategy will also be accelerated as the centrality degree of investors increases, but this effect is not significant compared with the other two players.

The investment volume of investors

To explore the sensitivity of three participants to the investment volume of investors, let P equal 3, 4, 5, 6, 7, and 10. The evolutionary trajectories of the three participants are shown in Figure 14. As seen from Figure 14A, P2P lending platforms are more willing to adopt the "positive disposal" strategy when the invest volumes are relatively low. However, when the invest volume increases gradually and exceeds the threshold, P2P lending platforms will change their strategy from 1 to 0. Similarly, Figure 14B shows that with the increase in the invest volume, there also exists a threshold for the investors, once the invest volume exceeds the threshold, the investors will change their strategies from 1 to 0. Figure 14C shows that with the increase in the invest volume, the evolutionary time of the financial regulators adopting the "co-governance regulation" strategy will be prolonged.

The initial cost of financial regulators to set up the co-governance regulation pattern

To explore the sensitivity of three participants to the initiative cost of financial regulators to set up the co-governance regulation pattern, let M_R equals 1, 2, 3, 4, 5, and 8. The evolutionary trajectories of the three participants are shown in Figure 15. When the initiative cost of financial regulators to set up the co-governance regulation pattern is low, the P2P lending platforms, the investors, and the financial regulators will

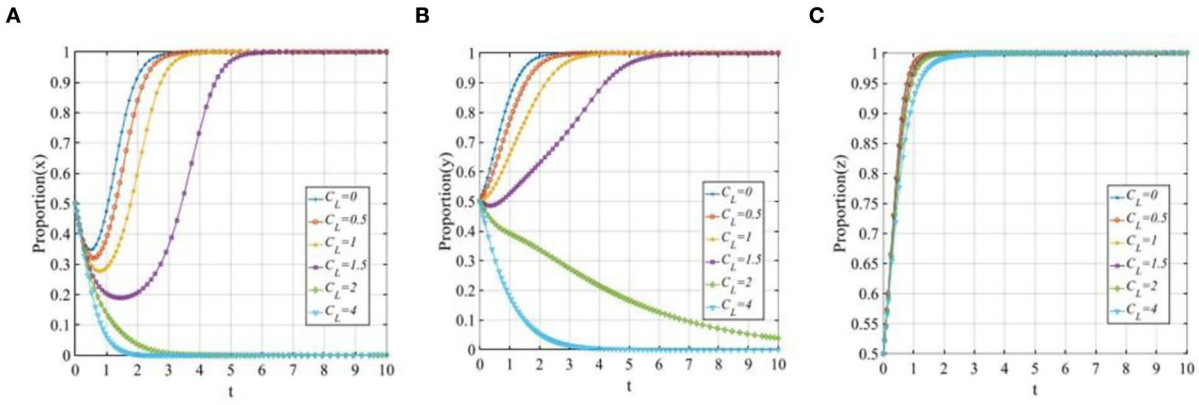


FIGURE 12
The sensitivity analysis of the cost of investors when $C_L = 0,\ 0.5,\ 1,\ 1.4,\ 2,\ 4$. **(A)** The sensitivity of P2P lending platforms; **(B)** The sensitivity of investors; **(C)** The sensitivity of financial regulators.

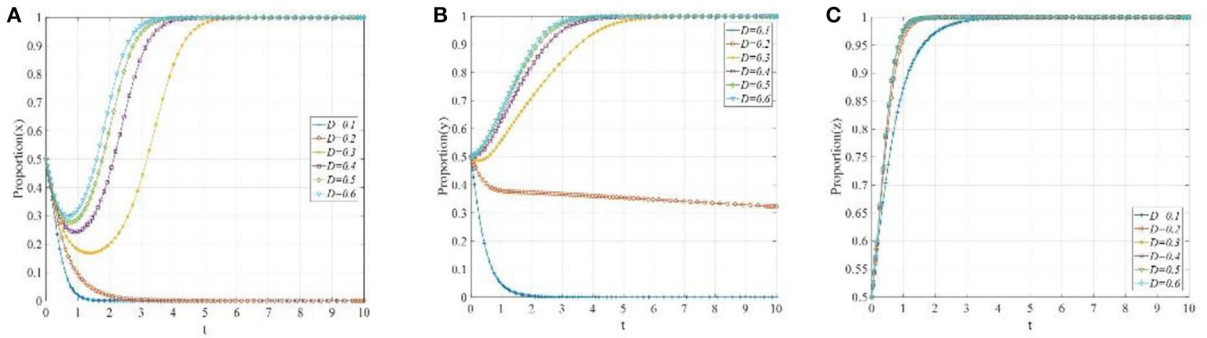


FIGURE 13
The sensitivity analysis of the centrality degree of investors when $D = 0.1,\ 0.2,\ 0.3,\ 0.4,\ 0.5,\ 0.6$. **(A)** The sensitivity of P2P lending platforms; **(B)** The sensitivity of investors; and **(C)** The sensitivity of financial regulators.

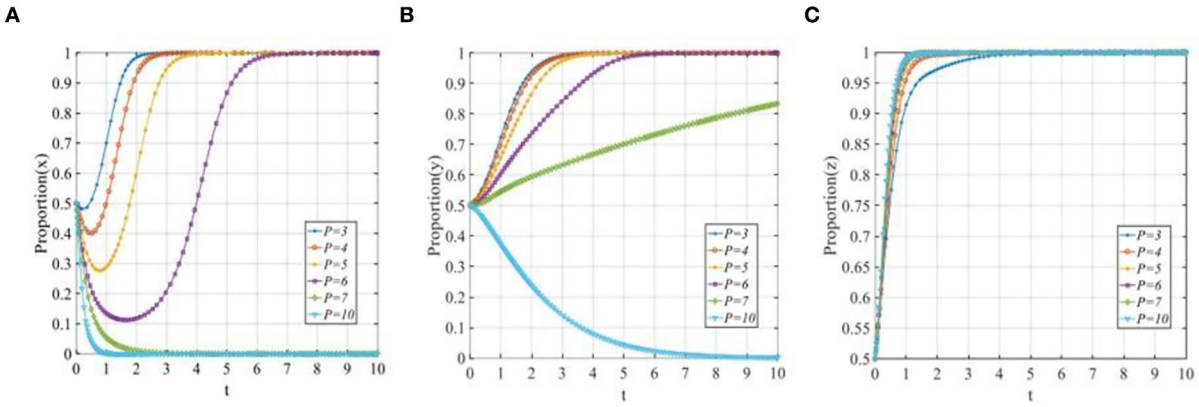


FIGURE 14
The sensitivity analysis of the investment volume of investors when $P = 3,\ 4,\ 5,\ 6,\ 7,\ 10$. **(A)** The sensitivity of P2P lending platforms; **(B)** The sensitivity of investors; and **(C)** The sensitivity of financial regulators.

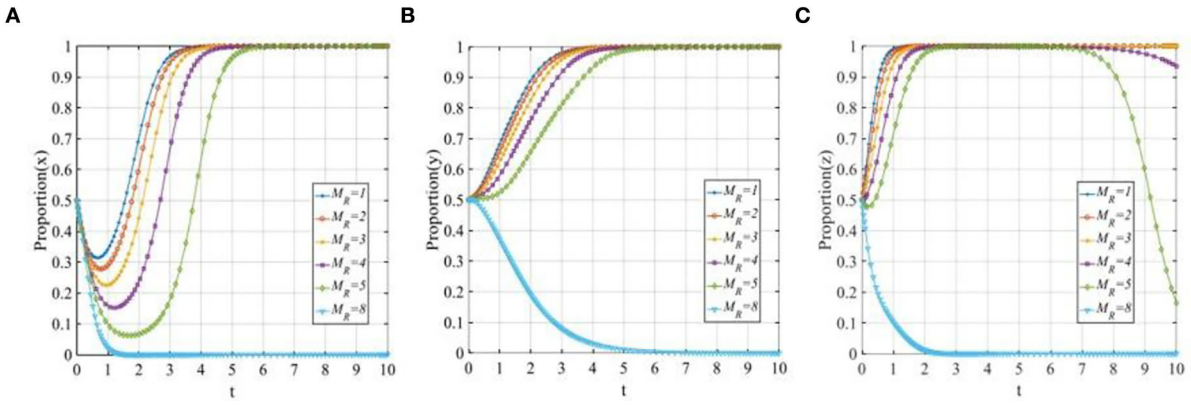


FIGURE 15
The sensitivity analysis of the initial cost of financial regulators to set up the co-governance regulation pattern when $M_R = 1, 2, 3, 4, 5, 8$. **(A)** The sensitivity of P2P lending platforms; **(B)** The sensitivity of investors; and **(C)** The sensitivity of financial regulators.

adopt the "positive disposal," "participating in co-governance," and "co-governance regulation" strategy, respectively. However, with the increase in the initial cost of financial regulators, their evolutionary time will be prolonged, once the initial cost of financial regulators is too high, i.e., $M_R = 8$, then all three participants will change their strategy from 1 to 0.

Discussion

Implications

From the theoretical analysis and numerical simulation of the tripartite evolutionary model, we find that several variables play important roles to enhance the final evolutionary stable state with the implementation of the social co-governance pattern.

The final evolutionary stable strategies are ultimately unchanged when the initial state changes and other parameters are fixed. Regardless of the initial ratio of strategies adopted by P2P lending platforms, investors, and financial regulators, the final ESS would be $E_8(1, 1, 1)$. This result indicates that the initial states of the three players are not related to the initial proportions, but their initial proportions would affect the evolutionary speed. The higher the proportions of P2P lending platforms adopting the "positive disposal" strategy, the investors adopting the "participating in co-governance" strategy, and the financial regulators adopting "co-governance regulation," the swifter the game would converge to $E_8(1, 1, 1)$. This reveals the necessity to build up the awareness of investors to participate in this co-governance pattern. In particular, the financial regulators should pay more attention to the P2P lending platforms with the low initial willingness of adopting "positive disposal" and enhance their supervision on those platforms. Above all, the financial regulators are expected to implement the co-governance regulation strategy to protect the rights and interests of the investors, in order to accelerate the P2P lending transformation and enhance the sustainable development of the FinTech industry. Specifically, we divide the overall evolutionary process into three stages according to the convergence states of the three players, as shown in Figure 16.

Stage I, the initial stage

At this stage, the proportion of financial regulators choosing "co-governance" presents an upward trend, while the proportion of investors choosing "co-governance" is on the rise and reaching a stable state. Since the initial stage of the establishment of a benign exit scheme of P2P lending platforms, financial regulators need to play the leading role of co-governance as much as possible to prevent the negative exit of the P2P lending platforms from damaging the rights of investors. However, due to the hysteresis of benign exit schemes and the legal of P2P lending platforms, the proportion of P2P lending platforms choosing "positive dispose" has gone through a process of first reducing and then increasing. At the same time, the enforcement and supervision of the benign exit of the P2P lending platforms are not mature enough, thus financial regulators should focus on improving the participation of investors in the benign exit process and establish regulatory norms to promote the benign exit process of the P2P lending platforms mainly by establishing regulatory norms.

Stage II, the transitional stage

The proportion that the P2P platform chooses to actively dispose of will rise to 1 as the proportion of the financial supervision department choosing co-governance supervision increases, and the proportion that the investor chooses to participate in co-governance is 1 and remains stable. At this

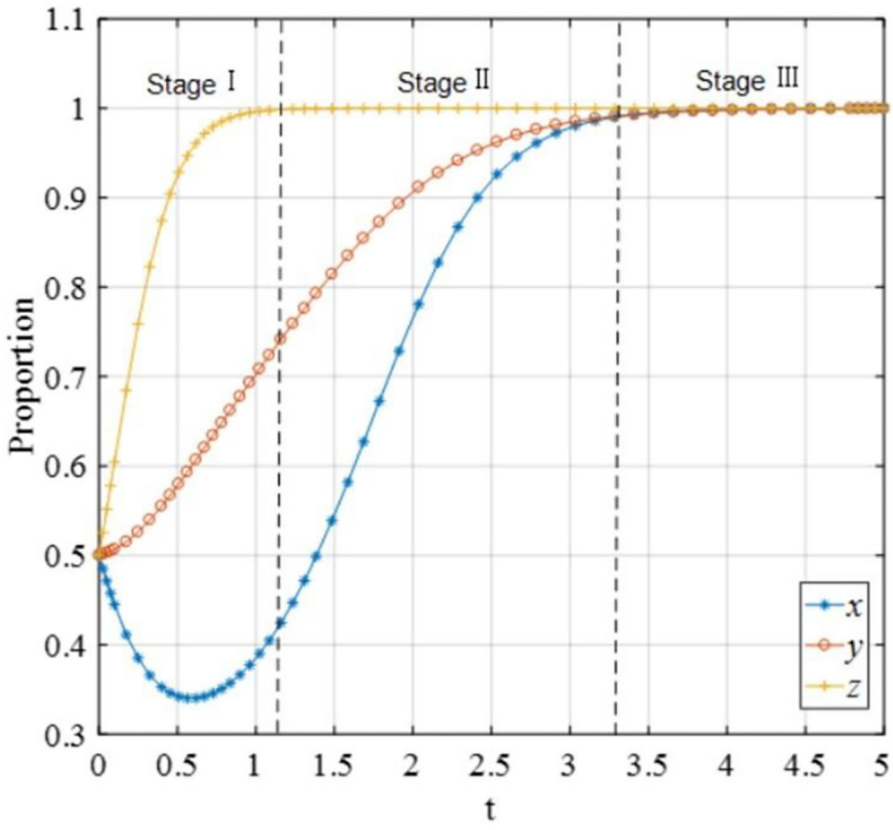


FIGURE 16
The evolutionary process with the co-governance scheme.

stage, since the P2P lending platforms received penalties from financial regulators and social reputation loss from investors, they will have the consciousness to support the benign exit process actively. However, the enforcement and supervision of the benign exit process is a long-term and arduous task, a co-governance pattern driven by financial regulators has been gradually formed in the benign exit of P2P lending platforms aimed at resolving conflicts of interest between the investors and P2P lending platforms. Thus, regulatory norms and value reconstruction will affect the behavior of stakeholders to participate in the co-governance in the benign exit of P2P lending platforms. To some extent, it demonstrates that co-governance of the financial regulators is still a guarantee for keeping P2P lending platforms to choose the strategy of "positive disposal."

Stage III, the mature stage

We can find the proportion of P2P lending platforms choosing "positive dispose," the probability of investors choosing "co-governance" and the probability of financial regulators choosing "co-governance" all remained stable at 1. At this stage, P2P lending platforms, investors, and financial regulators are ultimately stable and will not change their strategies due to the stable payoff, and the financial regulators can also avoid generating more sunk costs. The financial regulators and investors will collaborate to comply with the benign exit principles of P2P lending platforms, thus co-governance will become a by-product of the financial regulation. Various mechanisms for resolving conflicts of interest have been established at the level of the P2P lending industry to form an overall interest coordination mechanism for the benign exit of P2P lending platforms. Institutional norms and value reconstruction will affect the behavior of all stakeholders to participate in the co-governance of benign exit of P2P lending platforms. This has achieved the goal of benign exit of P2P lending platforms and formed a win–win situation for P2P lending platforms, investors, and financial regulators.

Applicability of our model

The outline of the regulatory document "Opinions on the classified dispose and risk prevention of P2P lending platforms"

regards the benign exit of P2P lending platforms as the top priority of regulation over P2P lending platforms and establishes a long-term mechanism for the FinTech loan industry. In fact, the co-governance of the benign exit of P2P lending platforms entails accelerating the exit and transformation of P2P lending platforms. This process will not only protect the legal rights and interests of investors but also drive one stream of P2P lending to be transferred into FinTech lending (Bussmann et al., 2020) in the emerging financial market. Thus, the necessary applicability and sustainability for the future development of the FinTech industry are provided.

However, at present, the benign exit process of P2P lending platforms is still faced with significant challenges. On the one hand, the benign exit lacks co-governance with the participation of investors, and it is excessively dependent on the regulation of financial regulators. On the other hand, the imperfection and hysteresis of the existing regulatory rules may lead to a lower level of efficiency in the benign exit process of P2P lending platforms, which would cause huge property losses to investors. In addition, in the process of benign exit, P2P lending platforms need to bear the high costs of asset disposals, while at the same time they are faced with difficulties in the operation process and often lack industry self-discipline. For investors, the lack of necessary funds to participate in the co-governance of benign exits of P2P lending platforms could lead to lower payoffs than expected. Therefore, all the stakeholders would have incentives to choose different strategies. For this reason, the strategies of the stakeholders are analyzed from the perspective of a tripartite evolutionary game, within which a truthful interpretation of the strategic choices is provided in the model results.

Conclusion

This study established a tripartite evolutionary game model for the first time to discuss the co-governance pattern in the benign exit process of P2P lending platforms in China. The main conclusions are as follows. (1) There are eight equilibrium points in the tripartite evolutionary game model, but only four ESSs. From the perspective of theoretical analysis and numerical simulation, the equilibrium point $E_8(1, 1, 1)$ is a more appropriate choice for the co-governance scheme of benign exit of P2P lending platforms, i.e., the one where P2P lending platforms adopt the "positive disposal" strategy, investors adopt the "participating in co-governance" strategy and financial regulators adopt the "co-regulation" strategy, respectively. (2) The initial proportion of P2P lending platforms adopting the "positive disposal" strategy, investors adopting the "participating in co-governance" strategy, and financial regulators adopting the "co-regulation" strategy cannot change the ultimate ESS, but will affect the evolutionary speed, the higher the initial proportion is, the faster the evolutionary speed will be. (3) Certain factors play a vital role in the strategic choices of P2P lending platforms, investors, and financial regulators. Specifically, the "positive disposal" proportion of P2P lending platforms increases with the increase in a penalty given by financial regulators to P2P lending platforms and the reputation losses of P2P lending platforms. While the "positive disposal" proportion of P2P lending platforms increases with the decrease in total costs of P2P lending platforms. Moreover, the threshold effect exists in the rewards given by financial regulators to investors, the cost of investors when adopting the "participating in co-governance" strategy, and the investment volume of investors, which may change the evolutionary stable strategies of P2P lending platforms, investors, and financial regulators.

Based on the above results, related policy strategies are proposed. (1) The financial regulators should adopt the co-governance pattern in the benign exit process of P2P lending platforms, to guarantee that P2P lending platforms and investors are able to obtain optimal benefits. This would help facilitate the benign exit process of P2P lending platforms and provide deeper insight into the mechanism of transformation in the FinTech industry. (2) The financial regulators should leverage regulation technology (RegTech) to reduce the initial costs of building the co-governance pattern and encourage investors to raise their awareness of participating in the benign exit process. (3) At the beginning of the benign exit of the P2P lending platforms, priority should be given to the willingness of P2P lending platforms that choose the "benign exit" strategy and investors who choose the "participating in co-governance" strategy. By doing so, we can effectively prevent capital losses and increase the probability of the benign exit of P2P platforms. (4) When P2P lending platforms choose a "benign exit" strategy, but the willingness of investors to choose "participation in co-governance" is low, the regulators can effectively address this problem by imposing higher penalties on P2P lending platforms and improving the incentives for investors.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

Ethical review and approval were not required for this study as the game models in this paper are based on hypotheses and numerical simulations, not real experiments. The "participants" in this paper are all simulated parameters, not patients/participants who actually participated.

Author contributions

QW and XL: conceptualization, methodology, and writing—review and editing. XL: software. QW: writing—original draft preparation. CZ: supervision and project administration. All authors contributed to the article and approved the submitted version.

Funding

This research was supported by the National Social Science Foundation of China (Major Project), Grant Number 17ZDA147 and the National Social Science Foundation of China (Key Project), Grant Number 14AZD033.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Abdullah, M., Mamat, M. P., Abang Morni, A., Kamri, T., and Fui, L. H. (2020). The economic impacts of rehabilitation of selabat mudflats nature reserve, Kuching, Sarawak. *Malaysian J. Geosci.* 4, 19–21. doi: 10.26480/mjg.01.2020.19.21

Ackerman, J. (2004). Co-governance for accountability: beyond "exit" and "voice". *World Dev.* 32, 447–463. doi: 10.1016/j.worlddev.2003.06.015

Agarwal, S., and Zhang, J. (2020). FinTech, lending and payment innovation: a review. *Asia-Pacific J. Fin. Stud.* 49, 353–367. doi: 10.1111/ajfs.12294

Aldasoro, I., Gambacorta, L., Giudici, P., and Leach, T. (2022). The drivers of cyber risk. *J. Fin. Stabil.* 60, 100989. doi: 10.1016/j.jfs.2022.100989

An, J., Di, H., and Liu, G. (2022). The influence of entrepreneurs' online popularity and interaction behaviors on individual investors' psychological perception: evidence from the peer-to-peer lending market. *Front. Psychol.* 13, 825478. doi: 10.3389/fpsyg.2022.825478

Bachmann, A., Becker, A., Buerckner, D., Hilker, M., Kock, F., Lehmann, M., et al. (2011). Online peer-to-peer lending-a literature review. *J. Internet Bank. Commerce* 16, 1.

Bartle, I., and Vass, P. (2005). *Self-Regulation and the Regulatory State, a Survey of Policy and Practice.* Bath: Centre for the Study of Regulated Industries; University of Bath School of Management. p. 24.

Birnbaum, S. (2016). Environmental co-governance, legitimacy, and the quest for compliance: when and why is stakeholder participation desirable? *J. Environ. Pol. Plan.* 2016, 1–18. doi: 10.1080/1523908X.2015.1077440

Brandl, B., and Hornuf, L. (2020). Where did FinTechs come from, and where do they go? The transformation of the financial industry in Germany after digitalization. *Front. Artif. Intellig.* 3, 8. doi: 10.3389/frai.2020.00008

Bussmann, N., Giudici, P., Marinelli, D., and Papenbrock, J. (2020). Explainable AI in FinTech risk management. *Front. Artif. Intellig.* 3, 26. doi: 10.3389/frai.2020.00026

Chen, H.-Y., and Tsai, C-h. (2017). Changing regulations of peer-to-peer lending in China. *Bank. Fin. Serv. Pol. Rep.* 36, 13–25.

Chen, J., and Li, W. (2019). Construction and perfection of the responsibility of third-party platform providers of the online food trade in China. *Biotechnol. Law Rep.* 38, 29138. doi: 10.1089/blr.2019.29138.wl

Chen, X., Chong, Z., Giudici, P., and Huang, B. (2022). Network centrality effects in peer to peer lending. *Phys. A* 2022, 127546. doi: 10.1016/j.physa.2022.127546

Chen, X., and Wu, L. (2019). Psychological capital in food safety social co-governance. *Front. Psychol.* 10, 1387-. doi: 10.3389/fpsyg.2019.01387

Chen, Y., and Sivakumar, V. (2021). Investigation of finance industry on risk awareness model and digital economic growth. *Ann. Operat. Res.* 7, 1–22. doi: 10.1007/s10479-021-04287-7

Cheng, H., and Guo, R. (2020). Risk preference of the investors and the risk of peer-to-peer lending platform. *Emerg. Market. Fin. Trade* 56, 1520–1531. doi: 10.1080/1540496X.2019.1574223

Cherednychenko, O. O. (2016). Cooperative or competitive? Private regulators and public supervisors in the post-crisis European financial services landscape. *Pol. Soc.* 35, 103–114. doi: 10.1016/j.polsoc.2015.12.005

Chunsheng, D. (2020). Three-party evolutionary game analysis of P2P network lending based on the theory of nonlinear system stability. *Chin. Manag. Sci.* 12, 1–13. doi: 10.16381/j.cnki.issn1003-207x.2019.1007

Davis, K. (2016). Peer-to-peer lending: structures, risks and regulation. *JASSA* 3, 37–44. doi: 10.3316/informit.419817919644101

Deng, J. (2022). The crowding-out effect of formal finance on the P2P lending market: an explanation for the failure of China's P2P lending industry. *Fin. Res. Lett.* 45, 102167. doi: 10.1016/j.frl.2021.102167

Eijlander, P. (2005). Possibilities and constraints in the use of self-regulation and co-regulation in legislative policy: experiences in the Netherlands-lessons te be learned for the EU? *J. Appl. Sci. Res.* 17, 899–902.

Emekter, R., Tu, Y., Jirasakuldech, B., and Lu, M. (2015). Evaluating credit risk and loan performance in online Peer-to-Peer (P2P) lending. *Appl. Econ.* 47, 54–70. doi: 10.1080/00036846.2014.962222

Feng, Y., Fan, X., and Yoon, Y. (2015). Lenders and borrowers' strategies in online peer-to-peer lending market: an empirical analysis of PPDAI.com. *J. Electr. Commerce Res.* 16, 242.

Freeman, L. C. (1978). Centrality in social networks conceptual clarification. *Soc. Netw.* 1, 215–239. doi: 10.1016/0378-8733(78)90021-7

Friedman, D. (1998). On economic applications of evolutionary game theory. *J. Evolution. Econ.* 8, 15–43. doi: 10.1007/s001910050054

Fu, X., Ouyang, T., Chen, J., and Luo, X. (2020). Listening to the investors: a novel framework for online lending default prediction using deep learning neural networks. *Inform. Process. Manag.* 57, 102–236. doi: 10.1016/j.ipm.2020.102236

Gelatt, J. P. (1992). *Managing Nonprofit Organizations in the 21st Century.* Phoenix, AZ: Oryx Press).

Giudici, P. (2018). FinTech risk management: a research challenge for artificial intelligence in finance. *Front. Artif. Intellig.* 1, 1. doi: 10.3389/frai.2018.00001

Giudici, P., Hadji-Misheva, B., and Spelta, A. (2019). Network based scoring models to improve credit risk management in peer to peer lending platforms. *Front. Artif. Intellig.* 2, 3. doi: 10.3389/frai.2019.00003

Giudici, P., Hadji-Misheva, B., and Spelta, A. (2020). Network based credit risk models. *Qual. Eng.* 32, 199–211. doi: 10.1080/08982112.2019.1655159

Gu, D., Ding, L., and Luo, P. (2018). Evolutionary game analysis on credit risk control in P2P online lending platforms. *R&D Manag.* 30, 12–21.

He, Q., and Li, X. (2021). The failure of Chinese peer-to-peer lending platforms: finance and politics. *J. Corp. Fin.* 66, 101852. doi: 10.1016/j.jcorpfin.2020.101852

Huang, R. H. (2018). Online P2P lending and regulatory responses in China: opportunities and challenges. *Eur. Bus. Org. Law Rev.* 19, 63–92. doi: 10.1007/s40804-018-0100-z

Iaione, C. (2016). The CO-city: sharing, collaborating, cooperating, and commoning in the city. *Am. J. Econ. Sociol.* 2016, 12145. doi: 10.1111/ajes.12145

Jin, Y., Cheng, C., and Zeng, H. (2020). Is evil rewarded with evil? The market penalty effect of corporate environmentally irresponsible events. *Bus. Strategy Environ.* 29, bse.2403. doi: 10.1002/bse.2403

Jo, H., and Harjoto, M. A. (2012). The causal effect of corporate governance on corporate social responsibility. *J. Bus. Ethics* 106, 53–72. doi: 10.1007/s10551-011-1052-1

Kim, K.-H., Kim, M., and Qian, C. (2018). Effects of corporate social responsibility on corporate financial performance: a competitive-action perspective. *J. Manag.* 44, 1097–1118. doi: 10.1177/0149206315602530

Ko, C. C., Liu, C. Y., Chen, Z. Y., and Zhou, J. (2019). Sustainable development economic strategy model for reducing carbon emission by using real options approach. *Sustainability* 11, 195498. doi: 10.3390/su11195498

Lee, E., and Lee, B. (2012). Herding behavior in online P2P lending: an empirical investigation. *Electr. Commerce Res. Appl.* 11, 495–503. doi: 10.1016/j.elerap.2012.02.001

Lei, X. T., Xu, Q. Y., and Jin, C. Z. (2022). Nature of property right and the motives for holding cash: empirical evidence from Chinese listed companies. *Manag. Decision Econ.* 43, 1482–1500. doi: 10.1002/mde.3469

Li, G., Elahi, E., and Zhao, L. (2022). FinTech, bank risk-taking, and risk-warning for commercial banks in the era of digital technology. *Front. Psychol.* 13, 934053. doi: 10.3389/fpsyg.2022.934053

Li, J., Hsu, S., Chen, Z., and Chen, Y. (2016). Risks of p2p lending platforms in china: modeling failure using a cox hazard model. *Chin. Econ.* 49, 161–172. doi: 10.1080/10971475.2016.1159904

Li, J., Xu, C., and Huang, L. (2022). Evolutionary game analysis of the social co-governance of E-commerce intellectual property protection. *Front. Psychol.* 13, 832743. doi: 10.3389/fpsyg.2022.832743

Lin, M., Prabhala, N. R., and Viswanathan, S. (2013). Judging borrowers by the company they keep: friendship networks and information asymmetry in online peer-to-peer lending. *Manag. Sci.* 59, 17–35. doi: 10.1287/mnsc.1120.1560

Liu, Q., Zou, L., Yang, X., and Tang, J. (2019). Survival or die: a survival analysis on peer-to-peer lending platforms in China. *Account. Fin.* 59, 2105–2131. doi: 10.1111/acfi.12513

Liu, W., and Xia, L.-Q. (2017). An evolutionary behavior forecasting model for online lenders and borrowers in peer-to-peer lending. *Asia-Pacific J. Operat. Res.* 34, 1740008. doi: 10.1142/S0217595917400085

Liu, Y., Zhou, Q., Zhao, X., and Wang, Y. (2018). Can listing information indicate borrower credit risk in online peer-to-peer lending? *Emerg. Market. Fin. Trade* 54, 2982–2994. doi: 10.1080/1540496X.2018.1427061

Luther, J. (2019). Twenty-first century financial regulation: P2P lending, FinTech, and the argument for a special purpose FinTech charter approach. *U. Pa. L. Rev.* 168, 1013.

Lyapunov, A. M. (1992). The general problem of the stability of motion. *Int. J. Control* 55, 531–534. doi: 10.1080/00207179208934253

Martinez, M. G., Fearne, A., Caswell, J. A., and Henson, S. (2007). Co-regulation as a possible model for food safety governance: opportunities for public–private partnerships. *Food Pol.* 7, 5. doi: 10.1016/j.foodpol.2006.07.005

Meng, C., Sun, L., Guo, X., Wu, M., Wang, Y., Yang, L., et al. (2021). Development and validation of a questionnaire on consumer psychological capital in food safety social co-governance. *Front. Psychol.* 11, 584810. doi: 10.3389/fpsyg.2020.584810

Offe, C. (1984). *Contradictions of the Welfare State*. Paris: Hutchinson.

Pan, W.-T., Zhuang, M.-E., Zhou, Y.-Y., and Yang, J.-J. (2021). Research on sustainable development and efficiency of China's E-Agriculture based on a data envelopment analysis-Malmquist model. *Technol. Forecast. Soc. Change* 162, 120298. doi: 10.1016/j.techfore.2020.120298

Peng, K., Wu, Z., Tang, J., and Haifeng, G. (2020). Research on the optimal structure state and supervision mode of P2P network lending market—based on the evolutionary game analysis of different supervision stages. *Syst. Eng. Theory Practice* 40, 2327–2338.

Rouvière, E., and Ca Swell, J. A. (2012). From punishment to prevention: a French case study of the introduction of co-regulation in enforcing food safety. *Food Pol.* 37, 246–254. doi: 10.1016/j.foodpol.2012.02.009

Smith, J., and Price, G. R. (1973). The logic of animal conflict. *Nature* 246, 15–18. doi: 10.1038/246015a0

Smith, J. M. (1988). *Did Darwin Get it Right?* Boston, MA: Springer.

Taylor, P. D. (1979). Evolutionarily stable strategies with two types of player. *J. Appl. Probabil.* 16, 76–83. doi: 10.2307/3213376

Tosun, J., Koos, S., and Shore, J. (2016). Co-governing common goods: interaction patterns of private and public actors. *Pol. Soc.* 35, 1–12. doi: 10.1016/j.polsoc.2016.01.002

Wainwright (1989). A dynamical systems approach to Bianchi cosmologies: orthogonal models of class A. *Class. Quant. Gravity* 6, 1409. doi: 10.1088/0264-9381/6/10/011

Wang, D., Zhu, W., Zhang, C., Li, H., and Wu, Y. (2020). Stakeholder symbiosis in the context of corporate social responsibility. *Symmetry* 12, 1897. doi: 10.3390/sym12111897

Wei, Z., and Lin, M. (2017). Market mechanisms in online peer-to-peer lending. *Manag. Sci.* 63, 4236–4257. doi: 10.1287/mnsc.2016.2531

Weng, Z., and Luo, P. (2021). The impact of guarantees on peer-to-peer lending platform: evolutionary game analysis and empirical evidence from China. *J. Theoret. Appl. Electr. Commerce Res.* 16, 2708–2731. doi: 10.3390/jtaer16070149

Wu, L., Liu, P., Lv, Y., Chen, X., and Tsai, F.-S. (2018). Social co-governance for food safety risks. *Sustainability* 10, 4246. doi: 10.3390/su10114246

Xia, P., Ni, Z., Zhu, X., and Ni, L. (2019). A novel key influencing factors selection approach of P2P lending investment risk. *Math. Probl. Eng.* 2019, 6086089. doi: 10.1155/2019/6086089

Xu, L., Zhou, Z., and Du, J. (2020). An evolutionary game model for the multi-agent co-governance of agricultural non-point source pollution control under intensive management pattern in China. *Int. J. Environ. Res. Public Health* 17, 72472. doi: 10.3390/ijerph17072472

Yan, L., Yin-He, S., Qian, Y., Zhi-Yu, S., Chun-Zi, W., and Zi-Yun, L. (2021). Method of reaching consensus on probability of food safety based on the integration of finite credible data on block chain. *IEEE Access* 9, 123764–123776. doi: 10.1109/ACCESS.2021.3108178

Yan, Y., Lv, Z., and Hu, B. (2018). Building investor trust in the P2P lending platform with a focus on Chinese P2P lending platforms. *Electr. Commerce Res.* 18, 203–224. doi: 10.1007/s10660-017-9255-x

Yang, D., Chen, P., Shi, F., and Wen, C. (2018). Internet finance: its uncertain legal foundations and the role of big data in its development. *Emerg. Market. Fin. Trade* 54, 721–732. doi: 10.1080/1540496X.2016.1278528

Yang, J., and Luo, D. (2017). The P2P risk assessment model based on the improved AdaBoost-SVM algorithm. *J. Fin. Risk Manag.* 6, 201–209. doi: 10.4236/jfrm.2017.62015

Yoon, Y., Li, Y., and Feng, Y. (2019). Factors affecting platform default risk in online peer-to-peer (P2P) lending business: an empirical study using Chinese online P2P platform data. *Electr. Commerce Res.* 19, 131–158. doi: 10.1007/s10660-018-9291-1

You, G., Guo, H., Dagestani, A. A., and Deng, S. (2021). Evolutionary game analysis of three players on behavioural strategy of P2P lending participants: a sustainable operations perspective. *Complexity* 2021, 7795016. doi: 10.1155/2021/7795016

Zhang, C., Liu, X., and Qi, W. (2020). The benign exit strategy of China's online lending institutions: an analysis based on evolutionary game. *Fin. For.* 25, 34–42.

Zhang, N., and Wang, W. (2019). Research on balance strategy of supervision and incentive of P2P lending platform. *Emerg. Market. Fin. Trade* 55, 3039–3057. doi: 10.1080/1540496X.2019.1624523

frontiers | Frontiers in Communication

TYPE Original Research
PUBLISHED 06 February 2023
DOI 10.3389/fcomm.2022.909427

Check for updates

OPEN ACCESS

EDITED BY
Oliver Niebuhr,
University of Southern Denmark, Denmark

REVIEWED BY
Claire Pillot-Loiseau,
UMR7018 Laboratoire de Phonétique et Phonologie (LPP), France
Juan Manuel Sosa,
Simon Fraser University, Canada

*CORRESPONDENCE
Sara Pearsell
✉ pearsels@mcmaster.ca

SPECIALTY SECTION
This article was submitted to Language Sciences, a section of the journal Frontiers in Communication

RECEIVED 31 March 2022
ACCEPTED 15 December 2022
PUBLISHED 06 February 2023

CITATION
Pearsell S and Pape D (2023) The effects of different voice qualities on the perceived personality of a speaker. *Front. Commun.* 7:909427. doi: 10.3389/fcomm.2022.909427

COPYRIGHT
© 2023 Pearsell and Pape. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

The effects of different voice qualities on the perceived personality of a speaker

Sara Pearsell* and Daniel Pape

Department of Linguistics and Languages, McMaster University, Hamilton, ON, Canada

Although previous studies investigated various aspects of voice quality perception and personality attribution there are no studies, to our knowledge, which simultaneously examine and compare the perception of various voice qualities when produced by the same individual. This work investigates how laryngeal and supralaryngeal voice quality variations of a speaker affect *listeners'* perceived personality traits (and thus perceived charisma) of that same speaker. Six Canadian English speakers produced paragraphs varying the following voice qualities: *modal*, *creaky*, *breathy* (natural and artificial), (hyper-)*nasalization*, and *smiling* (natural and extreme). Listeners of a perception experiment were then tasked to rate 10 statements for each presented audio stimulus. Statements were selected corresponding to a sub-section of the Big 5 personality traits shown to be linked to charisma perception. Results show significantly more positive listener ratings (i.e., higher ratings compared to *modal*) with medium effects sizes for both *smiling* variants across all personality traits. In contrast, *creaky* was perceived significantly more negatively overall for all personality traits, with a medium effect size. *Nasal* and *breathy* still achieved statistically significant rating differences compared to the *modal* baseline. However, the overall effect pattern was more complex, and effect sizes were small or negligible. Additionally, we found consistent differences for some voice qualities when examining listener ratings comparing male vs. female *speakers*: for both *creaky* and *smiling* (but not for other voice qualities), female speakers were rated more negatively when producing *creaky* for some personality traits, whereas both *smiling* variants were consistently rated higher for females compared to males.

KEYWORDS

variability in speech, perceived personality traits, voice quality variation, speech perception, laryngeal and supralaryngeal influences on cognition

1. Introduction

Recently there has been an increased interest in analyzing and understanding the effects of speakers' voice quality differences and how these *produced* differences may impact the *perception* of these speakers' personality traits. Areas of interest in this topic range from clinical techniques for best practices for a healthy vocal production (while avoiding vocal strain) to popular culture tips to sound more professional or speak more effectively. One specific area of interest is the role of voice quality in the perception of a speaker's personality traits, and, more generally, how these traits relate to the perception of speaker *charisma*.

1.1. Voice quality vs. vocal quality

Voice quality has been defined as "the quasi-permanent quality of a speaker's voice" (Abercrombie, 1967) and "those characteristics which are present more or less all the time that a person is talking. It is a quasi-permanent quality running through all the sound that issues from his mouth" (Abercrombie, 1967). Following this definition, and in line with researchers like Laver (1980) and Esling et al. (2019), voice quality differences are based on the specific auditory coloring of an individual's voice as a result of the variations of both *laryngeal* and *supralaryngeal* features which continuously occur throughout an individual's speech production. Several significant factors play a role for the variation of different laryngeal voice qualities: subglottal pressure (the air pressure below the vocal folds), medial compression (the contraction of the lateral cricoarytenoid muscles causing adduction; how tightly the vocal folds are pressed together), adductive tension (how tightly the arytenoid cartilages are pressed together at the posterior end of vocal folds) and longitudinal tension [the tension or slack of vocalis, thyroid and cricoid muscles, as well as the cricothyroid muscles (Laver, 1980)]. Following Laver's research, the most common phonation types, or laryngeal settings, are (i) *modal* or *normal voice*, the baseline (and non-pathological) voice setting, (ii) *breathy voice,* which has a high rate of air flow during production, (iii) *creaky voice* (also known as *vocal fry, laryngealization* or *glottalization*) characterized by very low frequencies which can be irregularly timed, (iv) *harsh voice*, a speech pattern with a normal fundamental frequency but aperiodicity or noise in the spectrum, and lastly (v) *tense or strained voice*, produced with a low rate of air-flow (often described as a "metallic" voice).

To conceptualize voice quality, it is helpful to think of each voice quality as a landmark on a continuum, with *breathy* on one end of that continuum (produced with a more open glottis), and *creaky* on the other end (produced with a constricted glottis). *Modal* voice is found between these two extremes. As a general notion of *modal* voice, this vocal quality has a more regular and periodic vibration pattern; there is no audible friction of the vocal folds, and the muscular tension is moderate. The vibrations are regular along of the vocal folds, often characterized as a "neutral mode of phonation" (Laver, 1980, p. 110). Medial compression, adductive tension, and airflow from the lungs are all moderate, and the longitudinal tension is low (vocal folds are shorter and thicker). The described voice quality landmarks vary slightly between individuals but maintain the same directional proximity to one another (*breathy* on one end, *modal* more central, and *creaky* on the other end).

Although these laryngeal features are the most dominant aspect in the description of different voice qualities, both the Abercrombie (1967) and Laver (1980) frameworks include suprasegmental modification of non-laryngeal features such as retroflexion/retraction, smiling or nasality. In this paper, we adapt these definitions of voice quality, which are also supported by the ANSI definition (i.e., that attribute of auditory sensation in terms of which a listener can judge that two sounds similarly presented and having the same loudness and pitch are dissimilar). Still, throughout this paper, we will use the term *vocal quality* to stress that this term would include both laryngeal features (e.g., modal, creaky, and breathy voice qualities) but also consistent and continuous suprasegmental feature variation (such as smiling and nasality, found on supralaryngeal and suprasegmental level). We hope that by using the term *vocal quality* we clearly define the inclusion of non-laryngeal vocal tract features since the term *voice quality* is often used very differently in the literature.[1] Please note that the suprasegmental features of interest in the current study, smiling and hypernasality, can of course co-occur with laryngeal voice quality variations. For the purposes of this study, the suprasegmental features of interest are produced with underlying modal voice quality, with modal voice representing the baseline measurement for each speaker.

1.2. Vocal quality and personality perception

Previous research has examined various acoustic features and perceptual cues and their relationship to personality trait attribution. Some of these studies have investigated the relationship between independent features of segmented speech signals such as f0 and pitch[2] (Puts et al., 2007; Rosenberg and Hirschberg, 2009; Quené et al., 2016; Berger et al., 2017), nasality and filled pauses (Möbius, 2003; Niebuhr and Fischer, 2019), amplitude or loudness (Novák-Tót et al., 2017), harmonics frequencies (Collins, 2000; Hodges-Simeon et al., 2010), and vocal quality (Wolf, 2015; Abdelli-Beruh et al., 2014) alongside their interaction with various personality traits. These results suggest that individual variation within physiological aspects of speech can play an important perceptual role in personality trait ascription.

In earlier research on the perception of vocal quality and perceived personality attributes, Pittam (1985) examined different vocal qualities of speakers and the impact of these qualities on listeners' ratings of solidarity, attractiveness, and status of the speaker. This study found that listeners' ratings of solidarity with a speaker were greater when there was the

1 Our decision to use vocal quality instead of voice quality for this paper stems from discussions with other researchers who often defined voice quality as purely consisting of laryngeal differences.

2 Pitch is a perceptual term taking into account the different acoustic properties of a complex acoustic waveform (normally consisting of the fundamental frequency and a number of optional harmonics). Therefore, pitch perception values can be different from measured fundamental frequency values.

presence of either *breathiness* or *whispery*[3] qualities in the speaker's voice. Perceptions of status were higher for *tense* voices as well as *breathy* voices compared to *whispery* and *nasal* voices (Pittam, 1985). In another study, Laver (1972) demonstrated an association of *breathy* voice with perceived higher sexuality and sensuality when the speaker was female but not when the speaker producing that *breathy* quality was male. Other studies have also demonstrated a correlation between certain vocal qualities and perceptions: the more significant the creakiness of a speaker, the higher the perceptions of that speaker's dominance or higher social status; the harsher the voice quality, the lower the perception of prestigious status (Esling, 1978; Scherer, 1979). Additionally, participants (who were described as young adults) rated voices with increased *creakiness*, above all the other vocal qualities assessed, as older. Esling (1978) and Scherer (1979) also suggest that this perception of age, as a result of the presence of *creakiness* in vocal quality, may account for the decrease in ratings associated with the friendliness and attractiveness of a speaker.

One major theory of personality and its associated traits is the *Big 5 of Personality Traits* (Norman, 1963; McCrae and John, 1992). Within this theory, personality traits are categorized and defined within five groups: Openness, Conscientiousness, Extroversion, Agreeableness, and Neuroticism (or OCEAN for short). The framework for our personality trait perception is based on this theory resulting in a broader categorization of attributes such as attractiveness or speaker status into one or more of these five personality traits.

Despite the interest in vocal quality and personality perception, to our knowledge, the current research remains limited to the focus of between-subject designs. That is, different voice quality conditions were always confounded with different speakers, for example, examining *creaky* vocal quality and thus only the influences of that *creaky* voice quality on listener perception. Understanding both the perception and production of multiple, potentially influential vocal qualities an individual is consistently able to produce can provide insight into many areas of interest. These interests could range from (1) understanding how listeners perceive a multitude of possible variations within a speaker's productions, as well as (2) clinical opportunities for those suffering from pathologies impacting their productions, (3) to professional opportunities for those who are outside academia to improve the effectiveness of their speech productions and understanding how their voice and its productions are perceived by an audience. Furthermore, currently it remains unclear what aspects of vocal quality variation are most salient for the concept of a *charismatic* speaker (Signorello and Demolin, 2013).

The present study investigates vocal qualities varied in a within-subject design, focusing on the following vocal qualities: *modal*, *breathy*, *creaky*, representing opposing ends of the voice quality spectrum as well as a medial point between the two, and the additional qualities of *nasality* (specifically hypernasality), and *smiling*. Within this study, these vocal qualities are rated explicitly in terms of within-subject personality traits and more specifically in terms of charisma-related traits.

1.2.1. Creaky voice

Creaky voice, also referred to as *vocal fry*, *glottal fry*, *laryngealization*, and *glottalization*, has been extensively researched. *Creaky* voice can be categorized by its irregular vocal fold vibrations created by the amalgamation of high adductive tension, low longitudinal tension, high medial compression and low subglottal pressure (Laver, 1980; Ladefoged and Johnson, 2011). It usually occurs at the lower end of a speaker's f0 range. Gick et al. (2013) explain that "In creaky voice, the vocal folds are very shortened and slackened to maximize their mass per unit length, and the IA (Inter Arytenoid) muscles are contracted to draw the arytenoid cartilages together. This action allows the vocal folds to stay together for a much longer part of the phonation cycle than in modal voicing..., only allowing a tiny burst of air to escape between long closure periods".

1.2.1.1. Creaky voice and personality perception

Previous research remains equivocal as to the perceptual influence of *creaky* voice on a speaker's personality characteristics. One study by Yuasa (2010) found favorable listener impressions for increased usage of *vocal fry*, with associations to personality traits such as professionalism, genuineness, and nonaggressiveness, as well as other positive assumptions about a speaker (e.g., higher level of education). *Creaky* voice has also been associated with worthiness, intelligence and friendliness (Pittam, 1987). However, other studies contradict these results: Anderson et al. (2014) showed that the presence of *creaky* voice, specifically in women, has the potential to negatively impact ratings of education level, competence, and trustworthiness. Gobl and Ní Chasaide (2003) found that *creaky* voice represents impressions linked to boredom and sadness. *Creaky* voice is found to be dominant in both younger male and female populations (Wolk et al., 2012; Abdelli-Beruh et al., 2014). Despite being present in both genders, research has shown when it comes perception of creaky voice, female speakers are more frequently perceived negatively compared to male speakers (Anderson et al., 2014; Wiener and Chartrand, 2014; Pointer et al., 2022). Although these studies present conflicting results, personality traits selected across studies do not equate to the same meaning or interpretation. It

3 Whispery voice is categorized as a combination of glottal friction and voicing. This combination creates greater amounts of inter-harmonic noise, creating an almost flat spectrum with increased levels of energy (Laver, 1980). Breathy voice differs from whispery because of weaker medial compression and a decreased degree of voicing effort. However, Laver (1980) notes the perceptual boundary is not clear between whispery and breathy. In this paper we use the term breathy quality.

should also be noted, regardless of personality trait mismatching across studies, that gender (and perhaps context) appears to influence the perceptual impact of *creaky* voice on listeners, therefore providing insights for the hypothesized outcomes of the current study when varying speaker and/or listener gender.

1.2.2. Breathy voice

In voices which are considered healthy (i.e., non-pathological), *breathiness* is categorized by partial adduction along the length of the vocal folds, with both the medial compression and adductive tension at low values, thus resulting in the increased escape of air (Laver, 1980; Reetz and Jongman, 2020). The amount of air escaping during phonation can cause differences in the perceived *breathiness* of a speaker's voice, with less adduction and a more gradual closing of vocal folds making the voice sound breathier (Hanson, 1997).

1.2.2.1. Breathy voice and personality perception

As previously described, *breathiness* has been shown to increase listeners' solidarity ratings[4] and perceived status (Pittam, 1985) as well as to influence perceived sexuality and sensuality for female speakers (Laver, 1980). However, research on the influence of this specific voice quality remains limited. Understanding the gap in the literature with respect to *breathy* vocal quality can provide further insight into how vocal qualities impact listeners' categorization of speakers' personality traits.

1.2.3. Nasal voice

Nasality is a vocal quality which results from nasal sound energy in the production of a speech signal. It is the result of the velopharyngeal port being either open or closed at inappropriate times or more than acceptable in a given language or dialect.

Nasal vocal quality is the acoustic result of the sustained and excessive coupling of the nasal and oral cavities during speech and can be categorized in one of two ways: hypernasality (i.e., going toward an excess of nasality) and hyponasality (i.e., going toward the absence of nasality). Hypernasality is caused by an excess of air leaking out through the nasal cavity when speaking. This results in extra (nasal) resonances in the acoustic speech stream. This type of nasality can be a result of several factors, from physiological issues, including structural problems (e.g., shortened soft palate or movement problem causing incomplete closure of the nasal cavity) to errors in sound acquisition (e.g., not learning, normally as a child, how to control the movement of air through the vocal tract cavities). Additionally, hypernasality still can have varying degrees of presence (more nasal and less nasal) and is primarily a result of both the size and status of the velopharyngeal port opening (Watterson and Emanuel, 1981; Warren et al., 1988); however, this is a separate factor from the presence or absence of nasality in speech production. Hyponasality is the opposite of hypernasality, in which not enough air can pass through the nasal cavity, resulting in a lack of nasal resonances in the speech signal as a result of a blockage or obstruction in the nasal cavity. This vocal quality is typical of the common cold (Tull, 1999).

1.2.3.1. Nasal voice and personality perception

To our knowledge, there is no previous research on the perceptual impact of *nasality* variation (specifically hypernasality) in non-pathological voices, presenting a knowledge gap in the literature on this vocal quality and its effect on speaker perception. It is important to note that, in principle, hypernasality could be combined with other vocal qualities, such as *breathy* or *creaky* voice. In our study, we restrict our examinations to the effect of *nasality* coupled with underlying *modal* voice, thus excluding combinations with other phonation types. Furthermore, nasal coupling is *continuously* produced by means of a lowered velum throughout the full duration of a sentence/paragraph production.

1.2.4. Smiling

The physiological movements involved in *smiling* include the widening of the mouth, retraction of the lips, the lowering of the tongue dorsum, and the tendency of a speaker to lower their jaw (Shor, 1978; Erickson et al., 2009). As a result of these movements, the vocal tract shortens, therefore altering the auditory perception of a speaker through an increase in formant frequencies as well as amplitude (Tartter, 1980). Tartter found that *smiling* has an audible effect on speech, generally associated with increased positive interpretations in a *smiling* condition.

1.2.4.1. Smiling and personality perception

A study by Vazire et al. (2009) explored the impact of the speaker's sex on the interpretation of listeners' *smiling* perceptions. The outcome of the study revealed two separate affective states, one for men and one for women. For women, *smiling* was viewed as a signal of trustworthiness and indicated warmth or enthusiasm to the listener. *Smiling* in men was interpreted as a lack of self-doubt, and increased confidence and calmness. Other research has found producing speech while smiling positively impacts speech perceptions, but has ceiling effects: excessive smiling does not increase the perception of charisma when compared to moderate smiling (Tschinse et al., 2022). For the present study, the inclusion of the *smiling* condition aims to reveal the connection between the effects of a *smiling* speaker on the perception of personality traits and effectiveness as a speaker when embedded in our experimental setup. Of particular interest for the current study, similar to the findings for *creaky* voice, is the mismatch in personality trait attribution when comparing (speaker) gender. Please note that

4 i.e., listeners' solidarity ratings with the perceived speaker, in other words the speaker currently being rated by the listener.

smiling, like hypernasality, could be combined with other voice qualities such as *breathy* or *creaky*. In our study, we will examine *smiling* only with an underlying *modal* voice.

These vocal qualities (modal, creaky, breathy, nasal, and smiling) have been examined individually and been ascribed personality trait correlates. As previously mentioned, there remains a lack of knowledge comparing these different vocal qualities, in combination, and across individual speakers. We hope to clarify the saliency of each of these vocal qualities when compared to each other, while simultaneously clarifying their interaction with respect to personality trait association.

1.3. Personality traits and charisma

The definition of charisma presented by Niebuhr and Fischer (2019) states: "*charisma is symbolic, emotional laden, and value-based communication style signaling leadership qualities such as commitment, confidence, and competence that affect followers' beliefs and behaviors in terms of motivation, inspiration, and trust.*" To further understand how to conceptualize charisma and charismatic speech research has looked at listeners' perceptual ratings for speakers' voices. These ratings were obtained through a series of presented statements correlating to charisma which listeners would rate from positive to negative, depending on the statement of each scale (Rosenberg and Hirschberg, 2009; Tskhay et al., 2018). For example, Rosenberg and Hirschberg (2009) found that *charismatic speakers* were associated with the (personality) traits of being enthusiastic, charming, persuasive, and convincing, all traits which can be found and categorized within the *Big 5* (John and Srivastava, 1999). As there is increasing interest in the sources of perceived charisma and more generally influential speakers, relying on vague interpretations of *charisma* is insufficient while using only the Big 5 of personality traits is too broad. By analyzing charisma within the traits of the Big 5 a clearer and more concrete interpretation of charisma can be established. The motivation behind our research is two-fold. Firstly, quantifying charisma based on the Big 5 allows for a targeted understanding of which attributes form different trait categories in order to create the concept of charisma, while concurrently allowing for a better "big picture" interpretation of personality traits perception using the Big 5.

Although *charisma* may not be a trait in and of itself, there are still many personality traits that coincide with charismatic features of speech, as noted in a paper by Michalsky and Niebuhr (2019). As the authors point out, studies by several other researchers have demonstrated the relationship of the *Big 5* traits to charismatic speech features. Antonakis et al. (2016) implemented a training program targeted to teach charisma to managers and business leaders using a system called *Charismatic Leadership Tactics (CLTs)*. The purpose of these CLTs was to make the concept of charisma more tangible to learners. Within their research, the authors demonstrate that confidence and self-assuredness are two facets which comprise charismatic speech. When examining these facets within the personality trait dimensions of the *Big 5*, these two facets fall into the *extroversion* personality trait (Costa and McCrae, 1992; John and Srivastava, 1999). Michalsky and Niebuhr (2019) also point out that the personality trait *agreeableness* relates to charismatic features, such as kindness, warmth, and development of trust while *conscientiousness* links to job performance and self-discipline (Costa and McCrae, 1992; John and Srivastava, 1999). Using just these examples, whether charisma is a personality trait in and of itself is debatable. Despite this, the traits associated with charismatic features of speech do have a relationship with personality traits and the *Big 5*, and exploration of charisma within the Big 5 traits could provide a more general concept and understanding of the interaction of charisma perception and the use of vocal quality production.

In order to determine how different vocal qualities are attributed to the perceived personality traits of a speaker as well as how personality traits relate to charisma, the concept of personality traits needs to be further defined. As briefly mentioned above, one prominent theory of personality dimensions is that of the *Big 5 of Personality Traits* (Norman, 1963; McCrae and John, 1992). In this theory, personality traits can be described and categorized into the following sets: Openness, Conscientiousness, Extroversion, Agreeableness, and Neuroticism (or OCEAN for short). It is important to note that each of these categories is a range of extremes. For example, extroversion is on one side of the spectrum while introversion is on the other (John and Srivastava, 1999). There is a scoring system which takes participant responses to a number of questions and rates these responses as a score from high (e.g., extroversion) to low (e.g., introversion). Figure 1 provides a visual representation and brief summary of each of the five main traits as well as the traits associated with high and low scores.

The first trait is openness. This is a personality trait tied to imagination and insight, as well as openness to new experiences. Individuals with higher ratings in this trait are often perceived as more creative and have a wide-ranging set of interests. They are open, artistic, curious, and imaginative. Individuals who rate low in this trait are resistant to new ideas, are unimaginative, dislike change, and do not like to try new things (John and Srivastava, 1999). Using the questionnaires within the studies by Rosenberg and Hirschberg (2009), and Tskhay et al. (2018), we manually classified each of the questions presented there into the *Big 5 framework*. This later became the structure for our experimental design. From these two studies, the particular trait of openness has not been strongly associated as an indicator of charisma. In our experimental design, we therefore opted to omit this particular trait.

The second trait, conscientiousness, is linked to a person's attention to detail, attentiveness, and goal-directed behavior. Those with a higher score in this trait are generally categorized

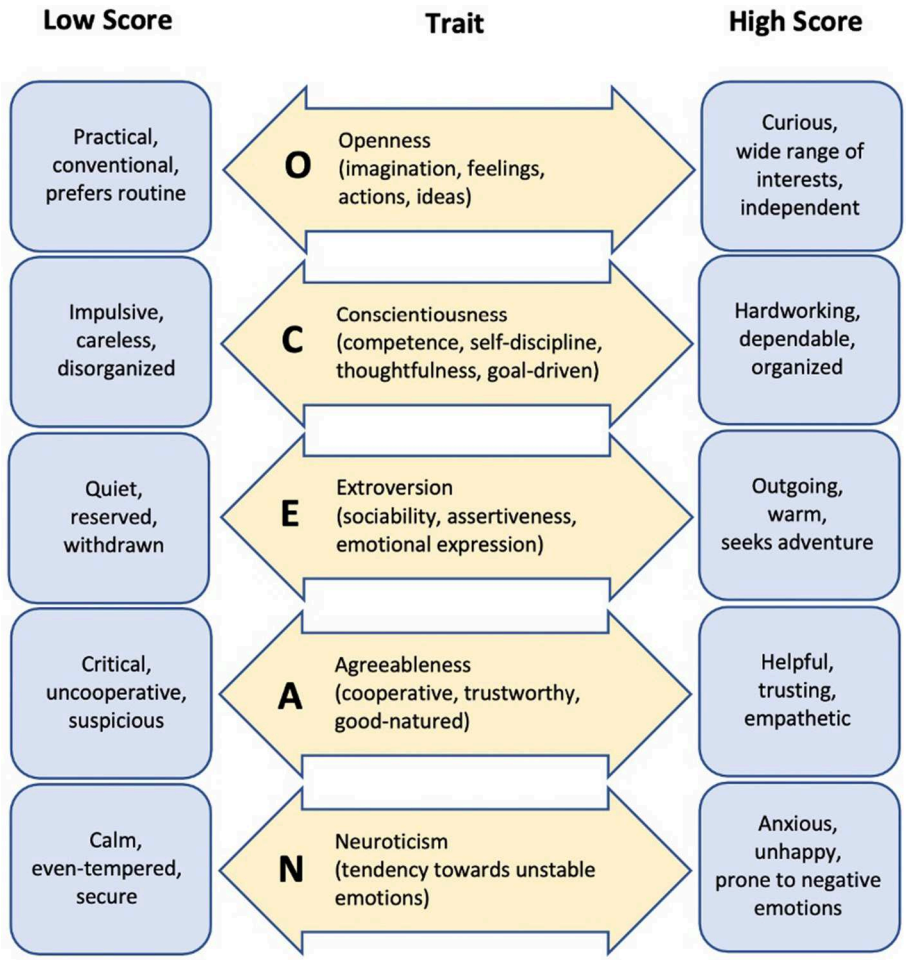


FIGURE 1
The Big Five personality traits, following Gray (2017).

as efficient, organized, reliable, and responsible, while lower scores are associated with those who are less organized and more flexible in their approach to work. They may also procrastinate, lack discipline, and be careless, resulting in difficulty in completing tasks or goals (John and Srivastava, 1999). From the questionnaire list by Rosenberg and Hirschberg (2009), higher scores in conscientiousness related statements correlated to charisma but were however less proportionate in the number of statements presented than traits like extroversion and agreeableness.

Extroversion, the third trait, is related to the level and degree to which a person seeks interaction with their environment focusing on the social component. Those rating high in extroversion tend to be more social, assertive, outgoing, talkative, etc., while introversion, or those on the low rank of this trait, tend to be more reflective, and reserved, preferring solitude, avoiding being the center of attention and tend to be fatigued by an excess of social interaction (John and Srivastava, 1999). Generally, a higher rating for extroversion is characteristic for charisma perception (Vergauwe et al., 2017), and extroversion is the Big 5 personality trait that receives the highest focus when determining charismatic attribution (Rosenberg and Hirschberg, 2009).

The fourth trait, agreeableness, determines how people treat their relationships with others. Unlike extroversion, agreeableness has to do with the pursuit of relationships with motivations concentrating on people's alignment and their interactions with others (John and Srivastava, 1999). Higher ratings in this trait indicate a person who is kind, forgiving, sympathetic, and trusting. Lower rating signal skepticism, stubbornness, a lack of sympathy, and a person who doesn't care about the feelings of others. The ability to connect with people as well as develop trust are just a few aspects which have also been demonstrative of charismatic speakers (John and Srivastava, 1999). Higher agreeableness scores appear to

also signal increased charisma in speakers (Rosenberg and Hirschberg, 2009).

The fifth and last trait is neuroticism. This is the trait which encompasses how an individual perceives the world, including the likelihood of inferring events as difficult or threatening as well as the inclination to experience negative emotions. People who rate high in this trait are anxious, tense, unstable, hostile, or irritable and experience dramatic shifts in mood. Those who rate lower are more emotionally stable, calmer, rarely feel sad or depressed, and do not often worry (John and Srivastava, 1999). In general, higher ratings of neuroticism have been shown to be negatively correlated with charisma and charismatic traits (Bono and Judge, 2004).

Although the Big 5 of Personality has traditionally been designed to be used by individuals based on introspection, the current study models questions/statements used in the previous studies of Rosenberg and Hirschberg (2009) and Tskhay et al. (2018). Within these studies, questions/statements were structured to be extrospective rather than the traditional introspective structure of Big 5 questionnaires. Other research (Hart and Hare, 1994; Ziegler et al., 2010) has demonstrated that ratings given by others fall closely within the range of ratings given from introspection. Theoretically, this means that results collected from our study's extrospective structure should produce data similar to those which would have been made by introspection.

1.4. Aims of the study and hypotheses

As previously described, although there are studies which investigate individual vocal quality perception and personality attribution, there are no studies, to our knowledge, that simultaneously examine the perceptual effects of various vocal quality changes produced by the same individual speaker on the perception of charismatic traits within the context of the *Big 5*. The current research investigates how vocal quality variation (breathy, creaky, nasal, and smiling) of different speakers affects listeners' perceived personality traits and thus charisma of these same speakers. We are also interested whether one of these voice qualities is most salient in high (positive) vs. low (negative) personality trait ratings by listeners. Furthermore, we want to examine the influence of gender on listener perceptions: here we are interested in both the influence of *speaker* gender on listener ratings, but also the influence of *listener* gender.

Apart from the differences between vocal quality categories, we are also interested to examine the effects of two within-category modifications for *smiling* and for *breathy* voice. With respect to smiling, following the research by Tschinse et al. (2022) we are interested to either replicate or dispute the observed ceiling effect for *normal* vs. *extreme* smiling condition with our within-subject design, all with respect to charisma ratings. With respect to breathy voice, we aim to introduce a technical, or more artificial, noise source modification in addition to the natural speaker-produced condition, thus examining the perceptual rating difference between a naturally produced breathy voice on the one hand vs. an artificially generated (technical) breathy voice on the other hand. The motivation here is to find out whether artificial noise added to the complete communication chain (and thus *not* modulated by laryngeal differences) would influence personality trait perception. In technical terms, the technical noise should be speech-shaped to make the conditions comparable and avoid adding another confound dimension.

We have the following hypotheses:

- H1: Lower listener perception scores, or negative ratings, for *creaky* voice across all speakers (resulting in a lower rating for all investigated personality traits, including neuroticism[5]), signaling a lack of perceived charisma in speakers. Lower scores for these traits in previous studies (Bono and Judge, 2004; Rosenberg and Hirschberg, 2009; Tskhay et al., 2018) have demonstrated a correlation to negative perceptions regarding speaker charisma.
- H2: Higher, or more positive, listener ratings in personality traits for *smiling*, with smiling having a positive correlation with speaker charisma.
- H3 (null hypothesis): Following Tschinse et al. (2022) we expect to see a ceiling effect for *smiling*, with the natural *smiling* productions expected to have almost identical rating scores compared to the extreme *smiling* condition.
- H4 (interactional hypothesis): We predict *speaker* gender to play a role in listener ratings. Specifically, female speakers will be rated more negatively when producing *creaky* voice (i.e., receive lower personality trait scores). For male speakers, we predict a less negative (or higher score) attribution of *creaky* voice compared to female speakers, thus bringing their ratings closer to *modal* voice ratings, meaning creaky voice for female speakers would be perceived less charismatically than their male counterparts.
- H5: For naturally produced vs. technical breathy voice we expect to see perceptual rating differences, with naturally produced breathy voice ratings lower for all examined personality traits and therefore rating lower in charisma. The reason for the lower expected ratings for natural breathy voice is that we assume that listeners are able to distinguish between noise as part of the speaker's laryngeal system (and thus being constantly modified by the speaker's production), whereas a channel-induced noise source could be better separated from the judged speaker characteristics,

5 Please see section Experimental setup for the explanation about neuroticism scores.

and thus would influence personality perception ratings less than the natural breathy condition.

2. Materials and methods

2.1. Stimuli

We selected two paragraphs consisting of multiple simple sentences as the basis of the acoustic recordings to be used in the perception study. The paragraphs were constructed to have a neutral valence to prevent any impact from positive or negative valence in listener interpretations of the voice. Each paragraph was ∼12 s long.

1) There is a house on the street and the kitchen door is open. Inside the kitchen, there's some table clothes in a basket. A spoon is on the table beside a coffee cup. I see a rug on the floor and magnets on the fridge.
2) The bedroom has two windows and a closet. A painting is hanging on the wall beside a clock. A dresser is across from the bed. Four drawers are in the dresser. There is a book and a lamp on the nightstand.

Six native Canadian English speakers (3 female, 3 male) recorded the paragraphs. Of these speakers, four were professional voice actors (2 male, 2 female), and two were Linguistics graduate students of McMaster University (1 female, 1 male). Due to lockdown restrictions associated with COVID-19, the four professional voice actors used their own high-quality microphones and adequate recording environments to record their productions and were directed and monitored *via* Zoom by the authors of this study. The two graduate student speakers recorded the stimuli using a high-quality microphone (Rode NT1A) and Focusrite Scarlet audio interface in the sound-proof booth of the Phonetics Lab at McMaster University. For all recordings, the microphone distance was specified to be around 10 cm, with the microphone being horizontally off-centre (from the lips) by ∼30–45 degrees.

To ensure that the speech stimuli sounded natural without artificial manipulation or distortions, each of the voice qualities of interest was naturally (speaker-) produced. Although it can be challenging to produce several different vocal qualities on cue, we assumed that professional voice actors, as well as graduate students in Linguistics, would be highly skilled in their ability to do so. To ensure all speakers were producing exemplary productions for all vocal qualities and would sound highly natural, we explained each vocal quality, then acoustically demonstrated the vocal quality, and then continuously directed speakers on how to produce it. This included producing the vocal quality continuously throughout the produced sentences (i.e., from the start of the production to the end of the production), a comfortable and natural speech rate (not too fast, or slow), as well as limiting pitch variation (as stable and flat f0 as possible), and amplitude variation (avoiding emphasis or stress). Once speakers were able to produce each vocal quality consistently and with the previously mentioned constraints (continuous vocal quality production throughout utterance, natural speech rate, stable and flat f0, stable and consistent amplitude distribution) over the given paragraphs, they were then recorded. Both paragraphs were repeated three times for all voice qualities. The best of these repetitions was then selected as the stimulus for the listeners (i.e., the repetition with the least variation in pitch, and amplitude, continuous vocal quality production throughout the utterance, and natural speech rate). Prosodic differences were as tightly controlled as possible across speakers and conditions through the continuous direction during practice and recording sessions and auditory checks of the stimuli by the researchers. However, prosodic characteristics like f0 or intensity differences were not artificially manipulated to avoid the introduction of artifacts and did possess some variation across speakers. Since we are examining these vocal qualities against the speaker's own modal production (in other terms the baseline) we hope that any differences in prosodic control and variation across speakers' production is less impactful than if comparing directly to other speaker's productions.

The voice qualities produced were *modal*, continuous *nasalization* (specifically hypernasalization; hypernasality was produced with a lowered velum from beginning of a sentence to its end), continuous glottalization (*creaky* voice), continuous *breathy* voice, and continuous *smiling* classified into two conditions: natural (where speakers were instructed to produce a natural, comfortable smile while recording stimuli; labeled SmilingN), and extreme *smiling* (speakers were instructed to smile excessively and to an extreme while producing the stimuli; labeled SmilingEX).[6] The smiling conditions were also visually monitored during recording sessions. Within the *breathy* voice condition, we included two distinct classes: a *natural breathy* voice production (as produced by the speaker, labeled BreathyN) and an *artificial breathy* voice production (labeled BreathyT). This *artificial breathy* production was created by taking the measurement of HNR (Harmonic to Noise Ratio) of each speaker's *natural breathy* production and overlaying a speech-frequency shaped noise signal onto their *modal* production with identical HNR measurement as the *natural breathy* production but with a rather technical (or speech transmission channel) noise overlaid.[7] In total, the stimuli consisted of 7 different vocal qualities, including a *modal voice* production for each speaker as the baseline.

6 Note that by "continuous" we mean produced from the start to the speech production to the end of the speech production.

7 The HNR measurement was done using Praat's algorithm using the object type "Sound: To Harmonicity: (cc)".

After recording, the audio samples were screened with the audio editor Amadeus Pro (Hairer, 2021) and carefully checked for achieved accuracy and consistency of each vocal quality production by the two authors of this study. Additionally, a steep high pass filter (80 Hz for male speakers; 150 Hz for female speakers) was applied to remove and attenuate any additional low-frequency noise which may have been a part of the original recordings.

The final stimuli count was 84 acoustic stimuli (*six speakers* x *two paragraphs* x *seven voice quality conditions*). Stimuli were not repeated, so each acoustic stimulus was only played once for each set of questions.

In the following, we present results for measuring the acoustic parameters of the produced stimuli in the three vocal qualities *modal*, *creaky* and *breathy* to confirm that all stimuli were produced consistently and according to the specifications outlined above. The acoustic measurements used were average speech rate, average fundamental frequency, its standard deviation, CPP, HNR, jitter and shimmer. These measurements are presented in Appendix A1. Generally, we found that both male and female speakers produce the stimuli in similar and expected ways. The average speech rate is approximately four syllables per second and does not vary systematically between the modal, creaky and breathy conditions. Furthermore, the average f0 and its standard deviation within speakers also remains consistent across conditions. Speakers show the expected decreases in creaky condition (Blomgren et al., 1998; exception: speaker EM) and very similar values for modal compared to breathy voice (except for speaker HK and to some extent MK). For breathy quality, decreased values for CPP (cepstral peak prominence) measurements are an indicator of breathiness in speech with smaller ratios representing greater differences in breathiness perception (Park et al., 2019; Murton et al., 2020). We avoided using HNR measurements as errors in location of individual pitch pulse onsets can strongly affect HNR (Hillenbrand, 1987). Since the parameter CPP strongly correlates with breathiness and is more resistant to errors in fundamental period location than HNR (Hillenbrand and Houde, 1996), we examined the CPP values and found all six speakers produce consistent differences between modal and breathy voice. With respect to creaky voice, the parameters jitter and shimmer are often used as acoustic correlates. Jitter and shimmer measure the acoustic irregularities of vocal fold vibration and are linked to roughness, hoarseness or breathiness of a voice with higher measurements correlating to increases to these aspects of speech (Blomgren et al., 1998). Specifically, jitter relates to frequency variation from cycle to cycle, while shimmer relates to amplitude variation (Murton et al., 2020). All of our speakers for almost all paragraph conditions demonstrate clear increases of jitter and shimmer in their creaky voice condition, corresponding to findings from Blomgren et al. (1998).

2.2. Participants

Twenty-seven participants took part in the perception study. They were primarily undergraduate students at McMaster University around the age of 20–23 (the majority of whom are studying Linguistics, Health Sciences or Psychology). All participants reported normal hearing and cognition. They answered a set of demographic questions including gender, age, acquired and spoken languages and musical education background. The experiment was conducted in a sound-proof booth at the Phonetics Lab at McMaster University using the Gorilla Experiment platform with wave file playback and using state-of-the-art acoustic playback conditions (Focusrite Scarlett audio interface, Sennheiser HD 598 linear frequency-response headphones). The duration of the experiment was around 60 minutes, including the pre-screening components.

2.3. Experimental setup

2.3.1. Scales

One effective way of eliciting the perception of personality traits is the use of continuous sliding scales. In voice quality and personality research, different researchers used very different types of scales. One study by Puts et al. (2007) examined the perception of dominance/authority through the use of scales to acquire ratings. The researchers posed questions to listeners about a speaker's voice, including the perception of a speaker's likelihood to win in a physical fight or the dominance or submissiveness of the speaker. Weiss and Moeller (2011) also utilized sliding scales to establish the likability of a speaker with the German antonyms *sympathisch—unsympathisch* (in English, a rough equivalent of *pleasant—unpleasant).* Several other studies (Rosenberg and Hirschberg, 2009; Berger et al., 2017; Niebuhr et al., 2018) have implemented statement-based questions, e.g., "The speaker is X," with a study-specific decision of which perceptual qualities are selected for X. Among these studies, there are variations with these statement-based questions; some are simply yes/no responses, while for others, responses are presented as a sliding scale from *strongly agree—strongly disagree*.

2.3.2. Presented statements

The statements used in the present study were based on research by Rosenberg and Hirschberg (2009) and Tskhay et al. (2018). In their research, Rosenberg and Hirschberg selected tokens based on their own judgement on whether they perceived the token as being either *charismatic* or *non-charismatic,* resulting in 26 stimuli with a mean length of 10.09s. The tokens were as context neutral as possible (e.g., "It's a pleasure to meet with you today."). For each of their 26 tokens, participants were asked directly to rate how charismatic the

sample was on a 5-point scale. Additionally, participants were then asked to rate additional 23 attributes, using statements based on previous literature on charisma (see below). Examining the Big 5 in relation to the questions presented by Rosenberg and Hirschberg and the research presented by Tskhay et al., openness was not a trait applicable to ratings of charisma and was therefore omitted in the current study. Due to the high number of vocal qualities in our study's design, 10 statements were presented rather than the original 26 of the Rosenberg and Hirschberg study to prevent an excessively long experiment, and these statements were first classified by personality trait type and then balanced according to the proportion of each personality type in Rosenberg and Hirschberg's study (five extroversion, three agreeableness, one conscientiousness, and one neuroticism). Modeling after previous research (Rosenberg and Hirschberg, 2009; Tskhay et al., 2018), the statements regarding the personality traits of agreeableness, extroversion and conscientiousness were designed to have higher scores of these personality traits corresponding to higher participant ratings, meaning ratings were more positively associated with that trait. For the neuroticism personality trait, the statement aimed to have lower scores for higher participant ratings. These lower scores have been positively correlated to charisma as high scores for this trait are often associated with more negative connotations such as anxiety and proneness to negative emotions; the higher the score for neuroticism, the more the trait is exhibited, the lower the score, the less the trait is exhibited. Despite all of our statements being framed positively [rather than both positively and negatively as in Rosenberg and Hirschberg (2009)], the results should not be skewed, as scores for the Big 5 relate to either high or low scores within each trait (openness, conscientiousness, extroversion, agreeableness, and neuroticism).

For the current study, these 10 statements were presented to the listeners. As described, these sentences were constructed by the researchers modeling the research of Rosenberg and Hirschberg (2009) and Tskhay et al. (2018) and created with a neutral valence to avoid any influence of positive or negative emotional connotations of the speech stimuli on listeners. The statements depicted the speakers as professors with the intention of establishing a relationship between the speaker and the (student) listener. Based off the previously mentioned results regarding the various vocal qualities, their uses, and the different speech environments or contexts in which they may be preferentially used, some of these vocal qualities within the current study might not be expected given the established context of an academic setting (i.e., that the speakers are "professors"). However, our rationale for labeling the speakers as professors was to prevent any other interpretations of social standing differences between different speakers, as well as between speakers and listeners, and speech environment (formal rather than colloquial like friends or family). All of these conditions could impact the perceptual ratings of speakers. Since many of the participants were university students at McMaster University, we decided it would be both interesting and relevant to characterize speakers as professors.

The ten statements were split in time, with the first screen containing the first 5 statements and the following screen containing the last five statements (Figure B1 in Appendix). This was done to prevent participants from being overwhelmed by excessive text content on one computer screen. Each screen played the audio stimuli once. Participants would slide the "button" to the desired location on each scale (for each statement) to represent how much they either agreed or disagreed with each statement presented (Figure B1 in Appendix B).

The 10 statements were related to four of the Big 5 of personality traits:

i) extroversion (5 statements like "This professor engages students in the classroom")
ii) agreeableness (3 statements like "This professor is positive and likable")
iii) conscientiousness (1 statement: "This professor is organized and detail oriented")
iv) neuroticism (1 statement: "This professor is convincing in the way they speak").

The listener's task was to judge the ten presented statements with respect to the simultaneously and acoustically presented audio file (and thus containing the different recorded vocal qualities). As described before, each of the audio files contained one paragraph and was recorded by the 6 individual speakers with seven different voice qualities. The statements on the screen (which were accompanied by the presented audio stimulus) would then be judged by the listeners using continuous sliding scales (from *strongly agree* to *strongly disagree)*. For our analyses, the *strongly disagree* end of the scale would be coded as 0% listener rating, and the *strongly agree* end of the scale would be coded as the 100% point of possible listener rating.

2.4. Statistical analysis

The statistical analysis was performed using the software R (R Studio Team, 2018) and RStudio (R Core Team, 2018). Parametric (e.g., ANOVA) or non-parametric (e.g., Wilcoxon) tests (depending on tests for normality of the data distributions) are used to determine whether listener responses/scores for each examined vocal quality (with respect to the presented statements and thus personality trait classification) as a dependent variable would be significantly different compared to the *modal* voice baseline (judging the exact same vocal quality stimulus and presented statement).

3. Results

3.1. Speaker-specificity vs. vocal quality influences

First, we aimed to examine whether each examined speaker would indeed drive a *vocal quality difference* in participant ratings or whether participants instead chose to rate an overall and general speaker personality (i.e., a personality *gestalt*) independent of the presented vocal quality manipulations and variations. In other words, we wanted to examine if listeners indeed showed an influence of varying vocal qualities *for each speaker* or rather judged each speaker based on his/her overall *vocal personality*.

To examine this question, we first present each speaker, and each examined vocal quality for listener ratings using violin plots with overlaid boxplots, as shown in Figure 2. Firstly, it can be seen that for each speaker, the median rating for *modal voice* differs, thus establishing an overall speaker effect on listener ratings. Additionally, certain speakers are judged more positively (i.e., achieve higher response ratings overall) than other speakers, e.g., speaker CS is rated more positively overall for all examined vocal qualities than, for example, speaker JF. Furthermore, the variation of vocal quality clearly shows an effect on listener ratings, with *creaky* voice obtaining consistently negative (lower) participant ratings (except for speaker EM) and *smiling* receiving consistently positive (higher) ratings (except for speaker SA). Please note, again, that for neuroticism, the scale is inversed as higher ratings for our scale correlate to negative neuroticism personality attributes. *Breathy* and *nasal* qualities show varying results compared to *modal voice* across different speakers, but it seems that their ratings rather closely correspond to the overall *modal voice* ratings for each speaker. Finally, there does not appear to be an influence from the speaker profession on listener ratings, as can be seen in Figure 2: professional voice actors (EM, CS, SA, HK) do not show apparent rating differences compared to the Linguistics graduate students (JF, MK).

Next, we present in Table 1 the correlation coefficients between *modal voice* and all other examined vocal qualities for each of the examined speakers and for all speakers combined. Theoretically, if listeners exclusively judge the acoustic personality of the underlying speaker (i.e., providing a rating of the speaker *gestalt* independent of the speaker-produced vocal quality), then correlations between modal voice and all other vocal qualities should be close to +1 (i.e., increasing *modal voice* ratings for that speaker should also increase all other vocal quality ratings simultaneously, thus excluding a possible effect of individual vocal quality on listener judgments). In contrast, if listeners exclusively judge the different vocal qualities (but choose to ignore the overall speaker identity), then correlations would strongly depend on the individual vocal quality comparisons. For example, it could be expected, based on the results in Figure 2, that the correlation between *modal* and *creaky* would be inversely related (i.e., closer to −1) compared to the correlation between *modal* and *smiling* (which could be positively correlated closer to +1), and all other comparisons showing varying correlations, but, most importantly, not being uniformly close to +1 as this would suggest an absence of a judged vocal quality difference. The correlation coefficients in Table 1 show that, for all 6 speakers, most vocal qualities obtain *varying* correlations (i.e., values *not* close to +1), thus establishing a clear influence of vocal quality on all listener ratings. For example, speaker SA shows a very high negative correlation between *modal* and *creaky* voice (i.e., if ratings for this speaker's modal increase, the ratings for creaky decrease), whereas this speaker's correlation between *modal* and *natural smiling* vocal quality condition is highly positively correlated (increased modal ratings correspond to increased *natural smiling* ratings), which clearly shows the influencing effect of *creaky* compared to *natural smiling* vocal quality on listener ratings. However, when examining the correlation table, it can also be shown that the vocal quality correlations are rather complicated and not straightforward (e.g., correlations between *creaky* and *modal* are highly positive for five speakers, and *smiling* vs. *modal* is highly negative for two speakers), but, importantly, the table, together with Figure 2, shows a clear influence of examined vocal qualities on overall listener ratings.

3.2. Vocal quality influences

To examine statistical differences between the examined manipulated vocal qualities, we first conducted a Shapiro-Wilk test for normality of the data distributions for each of the vocal qualities (i.e., one test for all *modal voice* responses, one test for all *nasal voice* responses and so on). All tests for normality were highly significant (see density plots of the seven vocal qualities in Figure C1 in Appendix C; see also the distributions of each violin in the violin plot in Figure 2), so we cannot assume a normal distribution of the data and thus decided to conduct significance tests using the non-parametric Wilcoxon signed-rank (matched sample) test. We performed pairwise comparisons to determine the statistical significance (1) of each vocal quality compared to the baseline *modal voice* and furthermore (2) comparing the *natural breathy* vs. *artificially breathy* and *normal smiling* vs. *extreme smiling* vocal qualities. In sum, 8 Wilcoxon tests were conducted, and the significance values shown in Table 2 are Bonferroni-corrected for these multiple comparisons. Effect sizes comparing each examined vocal quality compared to the modal voice perception are also reported. The table shows that *all* comparisons of the six examined different vocal qualities against the *modal voice* baseline are highly significant; thus, all 6 vocal qualities

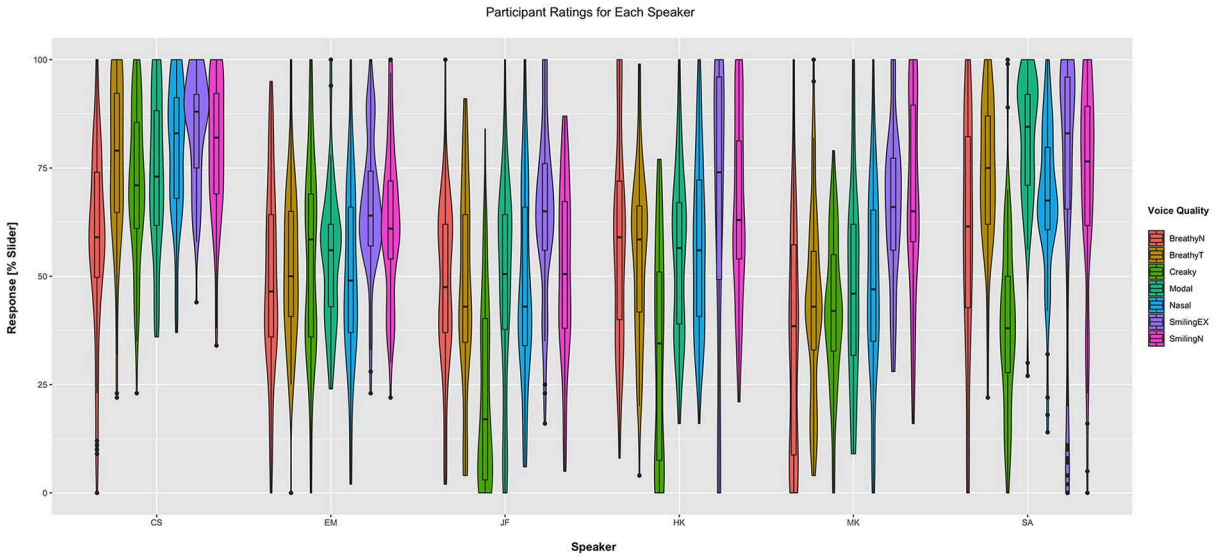


FIGURE 2
Voice quality ratings for all listeners, separately shown for each examined speaker (shown on the x-axis; the first three speakers are male, the last three female). The y-axis displays the slider position percentage (0% corresponding to the left extreme value of the slider and 100% to the opposite extreme). The colors represent the various examined vocal qualities. Shown are violin plots with overlaid boxplots.

TABLE 1 Correlation coefficients for *modal voice* vs. all other examined vocal qualities (i.e., correlating Modal-Creaky, Modal-Nasal, Modal-BreathyNatural, Modal-BreathyTechnical, Modal-SmilingNatural, Modal-SmilingExtreme), calculated separately for each speaker and combined for the six speakers.

Speaker	Gender	Creaky vs. modal	Nasal vs. modal	BreathyN vs. modal	BreathyT vs. modal	SmilingN vs. modal	SmilingEx vs. modal
CS	Male	0.97	−0.21	−0.78	0.94	0.09	−0.79
EM	Male	0.85	0.80	0.70	0.99	−0.75	−0.65
JF	Male	1.00	0.85	0.53	0.93	0.79	0.73
MK	Female	0.72	0.32	0.29	0.96	0.26	0.40
HK	Female	0.68	1.00	0.97	0.97	1.00	0.68
SA	Female	−0.90	0.86	0.90	0.39	0.90	0.83
All speakers		0.47	0.42	0.00	0.91	0.28	0.02

obtain significantly different listener ratings when compared to the perceived *modal voice* baseline. Comparisons of the effect sizes show a medium effect size for both creaky (rated lower or more negatively compared to modal voice) and the two smiling conditions (rated higher or more positive compared to modal voice; with extreme smiling having a higher effect size). In contrast, the natural breathy condition has a small effect size (rated lower or more negatively compared to modal voice), and all other vocal qualities have negligible effect sizes. Finally, the pairwise comparison of the two *smiling* conditions shows for the Wilcoxon test that they are perceived significantly different, and the same is true for the comparison of the two *breathy* conditions, which also shows highly significant differences.

3.3. Personality traits vs. vocal quality influences

In the following, we aim to examine the interaction between *vocal quality variation* and perceived *personality traits*. Figure 3 shows violin plots with overlaid box plots over all speakers, split by the four examined personality traits (x-axis) and examined vocal quality (colors). The baseline would be the rating of the perceived *modal voice*, and it can be seen that this vocal quality shows very similar values when comparing the four personality traits. Visual examination of the vocal quality differences for each personality trait confirms the results of the previously presented significance tests: mean listener ratings were higher, or more positive, for *smiling* for both the *natural smiling* condition

TABLE 2 Mean, median and standard deviations for each vocal quality (the bold-printed vocal quality values of the first column are reported) and results of the statistical *Wilcoxon signed-rank pairwise significance test* (z-values: third column; *p*-values: fourth column) comparing the two vocal qualities stated in column one, calculated over all participant responses.

	Median, mean and standard deviation	*z*-values (Wilcoxon test)	*p*-values (Wilcoxon test)	Effect sizes (comparison to modal)
Modal	56, 53.6, 24.3	–	–	–
Creaky vs. modal	38, 38.8, 26	−26.99125	$p < 0.001^{***}$	−0.61
Nasal vs. modal	52, 50.9, 25.6	−6.021475	$p < 0.001^{***}$	−0.11
BreathyN vs. modal	50, 47.6, 26.1	−12.03835	$p < 0.001^{***}$	−0.25
BreathyT vs. modal	54, 51.6, 24.2	−4.951314	$p < 0.001^{***}$	−0.08
SmilingN vs. modal	67, 65.8, 21.6	−25.61255	$p < 0.001^{***}$	0.50
SmilingEX vs. modal	70, 68.2, 22.2	−28.47034	$p < 0.001^{***}$	0.60
SmilingN vs. smilingEx	-	−7.67	$p < 0.001^{***}$	-
BreathyN vs. breathyT	-	−8.37	$p < 0.001^{***}$	-

All p-values are Bonferroni-corrected. The last column gives the effect sizes (Cohen's d) for each examined vocal quality compared to modal voice perception.
The *** symbol indicates statistically significant at the 0.1% level ($p < 0.001$).

and the *extreme smiling* condition across all personality traits. Again, the *extreme smiling* condition is rated higher compared to the *natural smiling* condition, and the natural breathy condition is rated lower than the artificial breathy one, corresponding to the significant differences observed in Table 2. Inversely, *creaky* was perceived lower, or more negatively, for all four personality traits. This is in line with previous research (Tartter, 1980) that general perceptions of *smiling* are correlated to more positive emotions and associations like trustworthiness, friendliness, etc., while *creaky* is perceived more negatively.

While Figures 2, 3 show the differences between the seven examined vocal qualities and thus gave an appropriate first overview of the obtained listener responses and their response distributions, the main aim of this study is to investigate the difference between an observed vocal quality and its corresponding *modal voice* perception, or, in other words, to see the pure effect of each vocal quality manipulation with respect to the four personality traits. In order to see this effect, we calculated, for each vocal quality judgement, the *difference percentage* between examined vocal quality and the baseline *modal voice* for that exact same acoustic stimulus comparison, thus effectively providing a pure effect of each vocal quality on listener ratings, split by personality trait. For example, we took the judgement of listener 1 judging the first paragraph of speaker 1 produced in *modal* voice and subtracted this value from the judgement of listener 1 judging the first paragraph of speaker 1 in a *creaky* voice, thus providing a measurement value showing the absolute difference in vocal quality rating (compared to *modal voice* judgements) for that specific speaker, listener, and paragraph identity. This calculation was then performed for all other (vocal quality, speaker, and listener) judgements. Thus, this difference quantifies the effect of the magnitude of change in vocal quality without taking into account other parameters. The results are presented in Figure 4, again as violin plots with overlaid boxplots. As can be seen, *smiling* again has the most considerable influence on all personality trait ratings, with more pronounced effects on agreeableness, conscientiousness and extroversion and a much smaller effect on neuroticism. In contrast, the *creaky* vocal quality has the strongest negative effect, with the most significant effect shown for conscientiousness compared to the other three traits. *Breathy voice* has a smaller negative effect on listener ratings, and interestingly this negative effect is strongest for neuroticism. An interesting result is the comparison of the two within-categories: the *natural smiling* vs. *extreme smiling,* and the *natural breathy* vs. *artificial breathy* condition. We do not observe the expected ceiling effect for *extreme smiling* conditions. Instead, for three of the four personality traits (excluding neuroticism), the *extreme smiling* condition consistently outperforms the *natural smiling* one, thus increasing the positive listener rating for more extreme smiling of each examined speaker. Interestingly, the *artificial breathy* condition generates more positive listener ratings compared to the *natural breathy* condition (see also Figures 2, 3 overall ratings), or, to turn it around, the *natural breathy* condition consistently leads to lower, or more negative, listener ratings compared to *artificial breathy* productions.

3.4. Effects of *speaker* and *listener* gender

When examining the effect of *speaker* gender, Figure 5 shows the mean differences in listener ratings, comparing listener judgements separately for male and female *speakers*, the produced vocal qualities and the four personality traits. We also provide results of the Wilcoxon pairwise significance

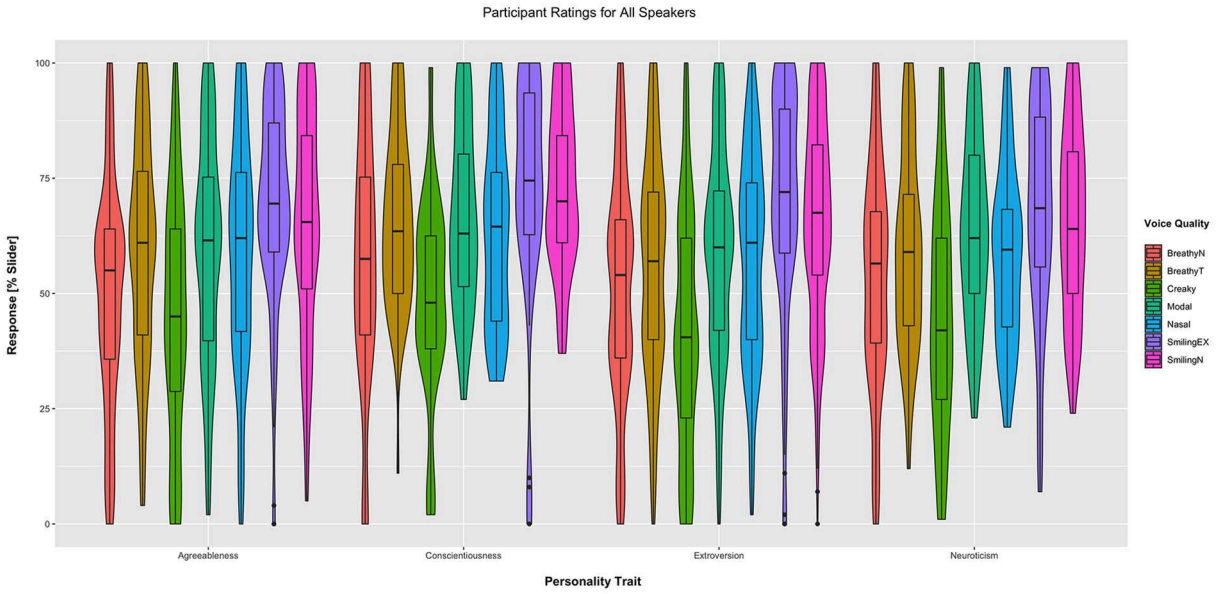


FIGURE 3
Voice quality ratings for all listeners aggregated over all six speakers' productions. The y-axis displays the slider position percentage (0% corresponding to the left extreme value of the slider and 100% to the opposite extreme). The x-axis shows the aggregation of the 10 statements into the four personality traits of interest: agreeableness, conscientiousness, extroversion, and neuroticism. The colors represent the various examined vocal qualities. Shown are violin plots with overlaid boxplots.

test in Table 3, over all speakers and for each examined vocal quality. Overall, the gender of the *speaker* has a significant effect on listener ratings ($p < 0.0001$), and all vocal qualities except *natural breathy* voice show a highly significant effect of speaker gender (see Table 3). Examination of the means in Figure 5 shows that there is a tendency that female speakers are judged more positively, independent of the examined personality trait and for almost all vocal qualities. Furthermore, we consistently see larger differences for two vocal qualities: *creaky* and, to some extent, *smiling*. Listeners rated female speakers more negatively when producing *creaky* voice compared to males for all personality traits. Our results thus confirm previous research (Anderson et al., 2014; Chao and Bursten, 2021) that demonstrated that *creaky* voice is frequently perceived negatively in women in a variety of environments. Additionally, both *smiling* variants are consistently rated higher for female speakers than their male counterparts.

Figure 6 shows mean plots comparing the effects of male and female *listener* gender on personality ratings, split by vocal quality and personality trait. Table 3 provides the significance results for each vocal quality. Overall, *listener* gender, similar to *speaker* gender, also has a significant effect on listener ratings ($p < 0.001$), however, the only vocal qualities rated significantly different when comparing the two (listener) genders are the *natural smiling* ($p < 0.001$) and *artificial breathy* ($p < 0.001$) vocal quality.

4. Discussion

Our results demonstrate that variation of the four different vocal qualities *breathy*, *creaky*, *nasal* and *smiling,* varied for the same individual speaker, can strongly and significantly influence listener perception of that speaker's personality traits, both positively and negatively: Results of the conducted Wilcoxon significance tests (see Table 2) show significantly higher listener rating scores for *smiling* voice qualities (for both *natural smiling* and *extreme smiling* condition) across all examined personality traits, whereas the *creaky* vocal quality was consistently and significantly rated lower for all personality traits for all participants, and thus perceived more negatively overall. For H1 we found that the continuous production of this creaky voice negatively impacts ratings of all personality traits. These results thus confirm the results of previous research regarding *creaky* voice and its unfavorable perception by listeners, which has linked this vocal quality to impressions of boredom and sadness (Gobl and Ní Chasaide, 2003), which would be classified into low neuroticism scores.[8] When specifically looking at *creaky* voice produced by women, the production of this vocal quality can negatively affect the ratings of competence (i.e., lower score in conscientiousness trait), trustworthiness (i.e., lower score in conscientiousness), and education level

8 Based on our inverted scale for neuroticism described in the Methods section.

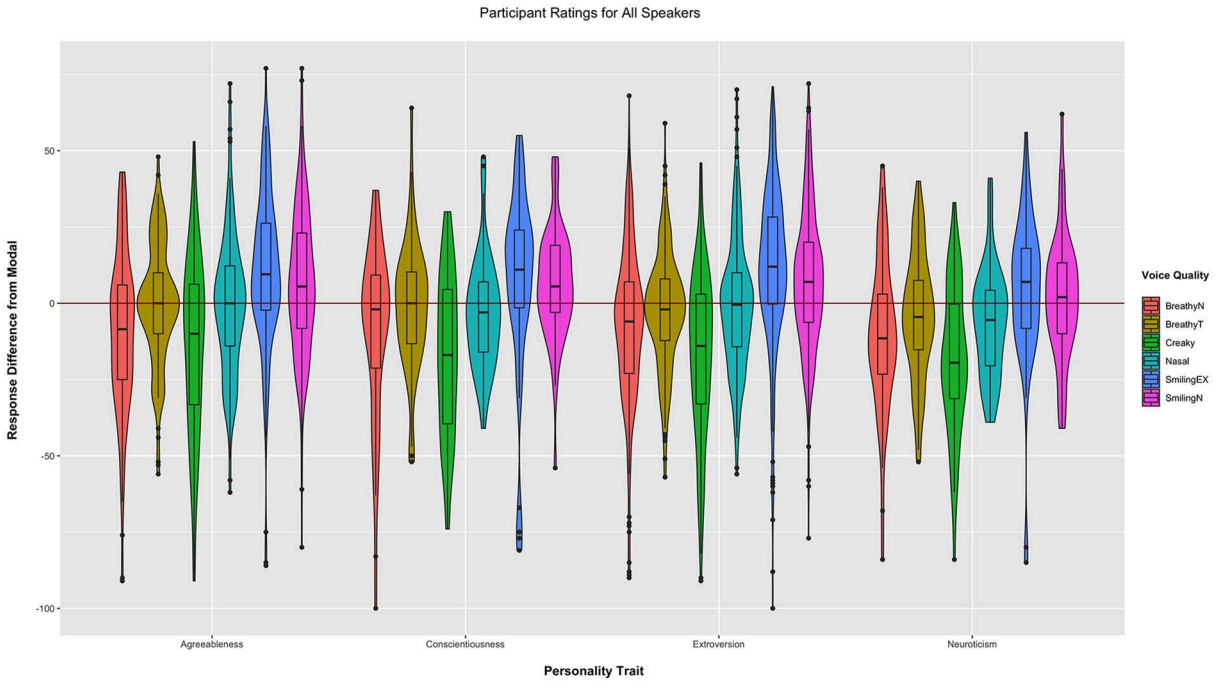


FIGURE 4
Violin plot and overlaid boxplots of the point-wise difference ratings for all listeners aggregated over all 6 speakers' productions. The y-axis displays the percentage difference as compared to each baseline (modal voice), i.e., the difference between each vocal quality rating and the corresponding modal voice quality (see text for further explanation). Positive rating difference percentages (compared to modal voice ratings) are above the 0% line; negative rating percentages are below this line. See Figure 3 for a description of axes and colors.

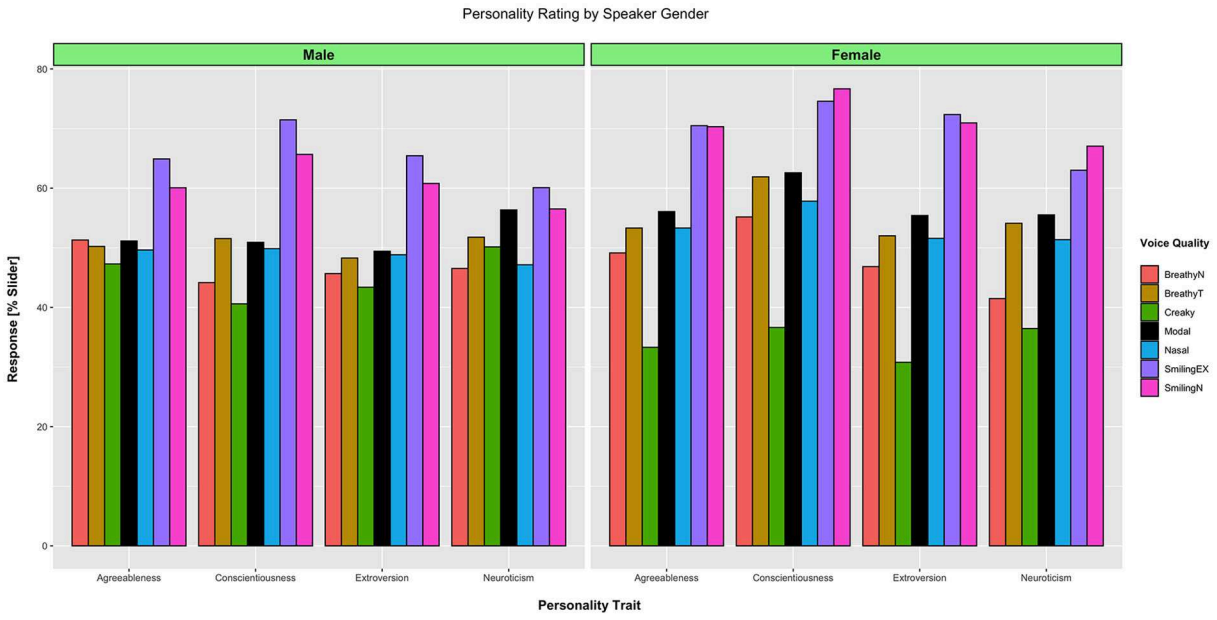


FIGURE 5
Mean ratings over all listeners comparing male *speakers* **(left)** vs. female *speakers* **(right)** productions, split by vocal quality. See Figure 3 for a description of axes and colors.

(Anderson et al., 2014). Furthermore, our results show this vocal quality is more negatively rated as compared to other vocal qualities such as *breathy* or *nasal* voice, both currently with very limited research results. Our results however are in contrast to those of both Yuasa (2010) and Pittam (1987), which found *creaky* voice correlated positively with professionalism, intelligence, friendliness, genuineness, and nonaggressiveness, as well as positive assumptions about a speaker, like assumed higher level of education. In sum, our study's lower ratings for all personality traits in combination with the results of previous studies (showing *decreased* scores for neuroticism, and conscientiousness being indicative of lack of charisma) all suggest that (continuously produced) creaky voice decreases the perception of speaker charisma. We can therefore accept this hypothesis.

For H2 our results are consistently and significantly higher ratings (see Table 2) for smiling. This confirms previous research, which found that *smiling* in women signals trustworthiness (high score in conscientiousness), indicated warmth (high score in extroversion trait) and enthusiasm (high score in extroversion trait) to the listener, whereas men who were *smiling* were interpreted as lacking self-doubt (high score in conscientiousness), confidence (high score in conscientiousness), and calmness (high score in neuroticism[9]; Vazire et al., 2009). This adds to general perceptions of *smiling* which are correlated to more positive emotions and associations like trustworthiness and friendliness (high score in extroversion; Tartter, 1980). Therefore, for H2 we can accept this hypothesis. Inversely to creaky voice, for *smiling* the observed higher ratings for all examined personality traits, in combination with the results of previous studies (showing *higher* ratings for neuroticism, extroversion, and conscientiousness being indicative of charisma), suggest that *smiling* positively impacts perceptions of speaker charisma.

For H3, also examining smiling, we find interesting results contrary to those presented by Tschinse et al. (2022). Our results in Table 2 show statistically significant differences between SmilingN (natural) and SmilingEX (extreme), with SmilingEX outperforming SmilingN with respect to positive listener ratings. These results thus reject our H3 null hypothesis. Although visually the differences between SmilingN (natural) and SmilingEX (extreme) appear rather small (see Figure 5), our statistical analysis demonstrates that increasing the smiling dimension also increases the positive influences on personality traits perception and therefore charisma. Whether the differences in results comparing our data with Tschinse et al. (2022) are due to our within-subject design or rather other methodological differences remains a cause for further study. Some of these mentioned methodological differences could be a result of stimuli: the stimuli in our study used short, isolated paragraphs (approximately 12 seconds) while the stimuli of Tschinse et al. (2022) were longer 1-min pitches. Prolonged auditory stimuli input allows for more habituation for participants and "saturation". Furthermore, the instruction for our extreme smiling and permanent smiling are not exclusively interchangeable, with the latter being temporally defined while the former is not.

TABLE 3 Results of the *Wilcoxon signed-rank pairwise significance test* (*z*-values and *p*-values) for each examined vocal quality and aggregated over all vocal qualities.

	Speaker gender		Listener gender	
	z-score	*p*-value	*z*-score	*p*-value
All (qualities)	−7.5408	<0.001***	−4.3274	<0.001***
Modal	−6.9697	<0.001***	−2.4725	0.01342
Creaky	−13.336	<0.001***	−2.1832	0.2367
Nasal	−4.976	<0.001***	−2.3980	0.01648
BreathyN	−1.1722	0.2391	−0.1885	0.8504
BreathyT	−5.4873	<0.001***	−3.74462	<0.001***
SmilingN	−14.0154	<0.001***	−3.40800	<0.001***
SmilingEx	−14.0154	<0.001***	−3.40800	0.2337

All p-values are Bonferroni-corrected (due to multiple comparisons).
The *** symbol indicates statistically significant at the 0.1% level ($p < 0.001$).

For the other two vocal qualities, *nasal* and *breathy,* no consistent and robust differences in listener ratings across speakers could be found. However, the Wilcoxon significance tests showed that these two qualities still obtained statistically significant differences, all compared to the *modal voice* baseline (see Table 2). Despite this fact, our results for both *breathiness* and *nasality* do not suggest a strong and clear trend relating these voice qualities to individual perceived personality trait differences since overall ratings of *nasality* and *breathiness* follow very similar trends as the *modal voice* baseline. Our results thus suggest that both nasality and breathiness do not play a salient role compared to smiling or creaky voice for personality trait attribution, although they both seem to lower listener scores for several traits for most speakers. For *nasality*, this result is quite interesting as it fills a current gap in the literature regarding the saliency and influence of *nasality* on personality trait perception and charisma. For *breathy* voice, previous research has suggested that this vocal quality influences perceptions of a speaker's perceived sexuality and sensuality, but only when the speaker is female (Laver, 1980). Also, solidarity perception with speakers is higher for *breathy* voices (Pittam, 1985). From our data, the results suggest that *breathiness* does not have that strong influence (but again, please note the significant difference to modal voice based on the Wilcoxon test).

Together, the results show increases or decreases in listener ratings for each vocal quality type. They appear to either *all* increase or decrease together, depending on the positive or negative perception of that vocal quality. By

9 Again, based on our inverted scale for neuroticism.

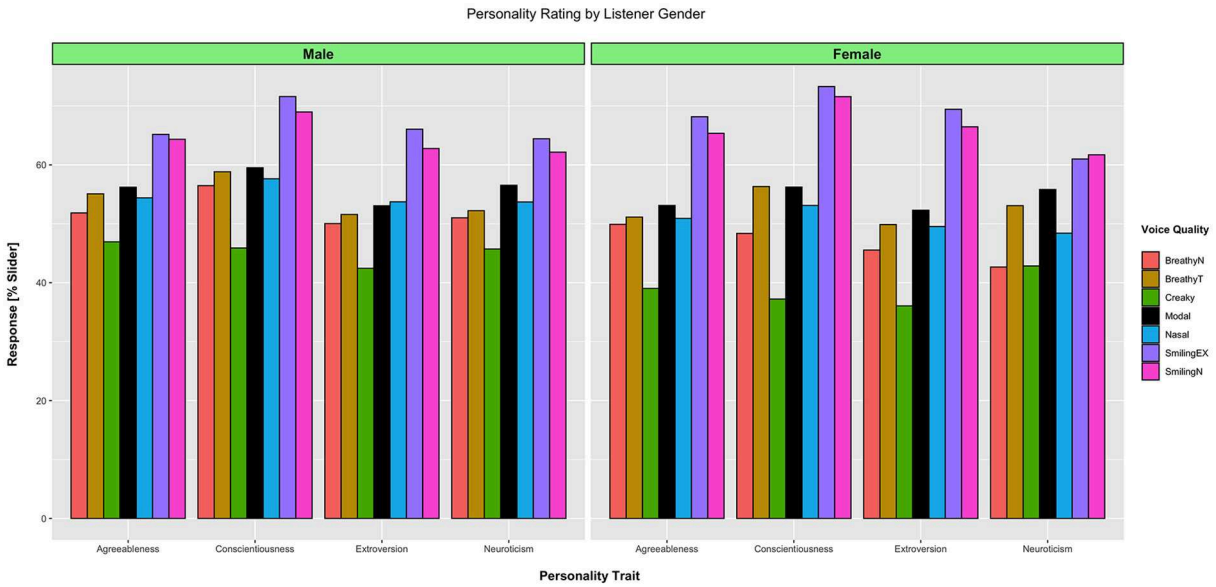


FIGURE 6
Mean ratings over all listeners comparing male *listeners* **(left)** vs. female *listeners* **(right)** productions split by vocal quality. See Figure 3 for a description of the axes and colors.

interpreting these traits collectively, we can see that those general increases/decreases of personality trait perceptions have a relationship with charisma; the higher, more positive, the ratings of traits, the greater the perception and saliency of charisma, whereas the lower, more negative the ratings, the lower the perception and saliency of charisma. Since we found that these increases/decreases in ratings synchronize across the different personality traits of the Big 5 (within each vocal quality), this can aid in future research on charisma in two ways: On the one hand, not all personality traits (of the Big 5) need to be utilized in experimental designs (i.e., only using questions/statements framed within the Extroversion trait, or Agreeableness, etc.) in order to capture meaningful interpretations of charisma and its presence. On the other hand, although statistically significant, some vocal qualities (nasal and breathy) are less salient in charismatic perception than other vocal qualities (creaky voice and smiling).

Further investigation into our within-category breathy voice differences reveal a statistically significant difference between BreathyN (natural breathy condition) and BreathyT (technical breathy condition). BreathyT is perceived with higher personality trait ratings and thus charisma, or, turning it around, BreathyN is perceived worse, thus confirming our H5. Despite being less salient than creaky vocal quality for charismatic trait attribution, these results suggest there is indeed a difference between adding the same type of noise (speech-shaped noise with identical HNR) to either the speaker's laryngeal signal (i.e., natural speaker-produced) or to the general communication channel (thus not being modulated by speech production differences). We speculate here that listeners are indeed able to separate the added channel noise from the speaker (personality trait) judgments, thus pointing to a hypothesis that added channel noise is not as detrimental to personality trait perception as noise directly produced by the speaker's larynx.

With respect to perceptual saliency and the magnitude of participant rating differences, certain vocal qualities are more pronounced than others for personality attribution. For example, see the difference in point-by-point comparison of *creaky* vocal quality (compared to *modal*) vs. *nasal* vocal quality (also compared to *modal*) as shown in Figure 4. Although previous literature has shown the various impacts of how speakers are judged by listeners regarding vocal quality differences, the current results can provide a better understanding which vocal qualities may require more focus when attempting to increase charisma perception: *smiling* more and avoiding continuous *creaky* voice appear to be more relevant and more salient than avoiding *nasal* or *breathy* productions.

Furthermore, we examined the effect of speaker- and listener-dependent factors, namely the effect of gender for both *speaker* and *listener*. For H4, rating differences comparing male and female *speakers* and their vocal quality show consistent differences for both *creaky* and *smiling* vocal quality but not to the same extent for the other vocal qualities. Listeners rated female speakers more negatively in *creaky* voice than the corresponding male speakers for the personality traits of agreeableness, extroversion, and neuroticism. Here, our results confirm previous research (Anderson et al., 2014; Chao and Bursten, 2021) that demonstrated that *creaky* voice is frequently

perceived negatively in women in a variety of environments. Additionally, both *smiling* variants are consistently rated higher for female speakers than their male counterparts. In sum, our results confirm that gender strongly influences the perception of vocal quality, both overall and within different personality trait contexts and we can accept this hypthothesis.

One limitation of the present study is the relatively small number of speakers and the observed inter-speaker variation (see, e.g., Figure 2, where listener ratings for speaker EM show the opposite pattern for *creaky* vs. *modal* voice compared to the other five speakers). Six speakers can provide a general picture of vocal quality and personality attribution, but this picture is still limited in the scope of potential variation, which may naturally occur in the production of individual speech patterns. Also, as can be seen in the correlation table, vocal qualities across speakers are not judged uniformly, thus introducing speaker-specific variation in this vocal quality study. For future studies, a higher number of different speakers could provide a more detailed understanding of the effect of vocal quality variations on charismatic traits, and it might continue to examine the more fine-tuned effects of *nasality* and *breathy* voice (in both technical and natural variation) that gave significant overall differences compared to *modal voice* perceptions but failed to provide a clear trend of the effects on individual personality traits.

An additional point of limitation in the current study is that we could not control for several other factors of variability for perceived vocal quality and personality ratings. For example, different possible settings would feed into the concept of *charismatic* speech and influence the ratings, for example: different communication contexts (formal vs. informal), environmental settings (e.g., academic, as in this study, vs. peer ratings), types of audiences (e.g., interviewers vs. colleagues), and of course whether the purpose of the communication is to be persuasive. Specifically, the established speaker-listener relationship in our study (i.e., the speaker being defined as a *professor* for our student participants) could influence vocal qualities to be perceived differently than if that relationship would have been established with a different social relationship paradigm (e.g., the listener is not a student, the listener is rating a friend, the listener is rating a co-worker, etc.). Since we chose this university setting—as previously explained—there is of course the chance that some of these voice qualities could be perceived differently for different speaker settings (friends or family) or different social environments (e.g., work, socializing with friends).

Finally, it is essential to note that other factors in combination with vocal quality appear to play a role, such as age vs. creakiness (Esling, 1978; Scherer, 1979) or the interaction of creakiness, f0, and speech rate (Parker and Borrie, 2018). These, along with many other variations, suggest that, of course, vocal quality is not the only important component for listeners when giving ratings of personality traits. For future studies, the inclusion of other speech features like f0 variation, speech rate differences etc., could provide a further understanding of the interactions between linguistic speech variation and voice quality.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary material, further inquiries can be directed to the corresponding author.

Ethics statement

The studies involving human participants were reviewed and approved by McMaster Research Ethics Board. The patients/participants provided their written informed consent to participate in this study.

Author contributions

SP and DP designed the experiment and stimuli, analyzed and interpreted data, and drafted and developed the manuscript. SP recorded the stimuli and collected participant data. DP also supervised the experiment. Both authors contributed to the article and approved the submitted version.

Funding

This work was supported by the NSERC Discovery Grant PGGPIN-2018-06518 to DP.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcomm.2022.909427/full#supplementary-material

References

Abdelli-Beruh, N. B., Wolk, L., and Slavin, D. (2014). Prevalence of vocal fry in young adult male american english speakers. *J. Voice.* 28, 185–190. doi: 10.1016/j.jvoice.2013.08.011

Abercrombie, D. (1967). *Elements of General Phonetics*. Edinburgh: Edinburgh University Press.

Anderson, R., Klofstad, C., Mayew, W., and Venkatachalam, M. (2014). Vocal fry may undermine the success of young women in the labor market. *PLoS ONE.* 9, e97506. doi: 10.1371/journal.pone.0097506

Antonakis, J., Bastardoz, N., and Jacquart, P. (2016). Charisma: an ill-defined and ill-measured gift. *Annu. Rev. Organ. Psychol. Organ. Behav.* 3, 293–319. doi: 10.1146/annurev-orgpsych-041015-062305

Berger, S., Niebuhr, O., and Peters, B. (2017). "Winning over an audience - A perception-based analysis of prosodic features of charismatic speech," in *Proceedings of the 43rd Annual Conference of The German Acoustical Society, Vol. 43* (DAGA), 1454–1457.

Blomgren, M., Chen, Y., Ng, M. L., and Gilbert, H. R. (1998). Acoustic, aerodynamic, physiologic, and perceptual properties of modal and vocal fry registers. *J. Acoust. Soc. Am.* 103, 2649–2658. doi: 10.1121/1.422785

Bono, J. E., and Judge, T. A. (2004). Personality and transformational and transactional leadership: a meta-analysis. *J. Appl. Psychol.* 89, 901. doi: 10.1037/0021-9010.89.5.901

Chao, M., and Bursten, J. R. (2021). Girl talk: understanding negative reactions to female vocal fry. *Hypatia.* 36, 42–59. doi: 10.1017/hyp.2020.55

Collins, S. A. (2000). Men's voices and women's choices. *Anim. Behav.* 60, 773–780. doi: 10.1006/anbe.2000.1523

Costa, P. T., and McCrae, R. R. (1992). *NEO PI-R Professional Manual.* Odessa, FL: Psychological Assessment Resources.

Erickson, D., Menezes, C., and Sakakibara, K. I. (2009). "Are you laughing, smiling or crying?," in *Proceedings: APSIPA ASC 2009: Asia-Pacific Signal and Information Processing Association, 2009 Annual Summit and Conference* (Asia-Pacific Signal and Information Processing Association), 529–537.

Esling, J., Moisik, S., Benner, A., and Crevier-Buchman, L. (2019). *Voice Quality The Laryngeal Articulator Model (No. 1)*. Cambridge, United Kingdom: Cambridge University Press. doi: 10.1017/9781108696555

Esling, J. H. (1978). *Voice Quality in Edinburgh: A Sociolinguistic and Phonetic Study. Ph.D. Dissertation.* Edinburgh United Kingdom: The University of Edinburgh.

Gick, B., Wilson, I., and Derrick, D. (2013). 'Articulating Laryngeal Sounds' in *Articulatory phonetics*. New York: John Wiley and Sons. p. 109.

Gobl, C., and Ní Chasaide, A. N. (2003). The role of voice quality in communicating emotion, mood and attitude. *Speech Commun.* 40, 189–212. doi: 10.1016/S0167-6393(02)00082-1

Gray, R. (2017). *Boundless Psychology*. Lumen. Available online at: https://courses.lumenlearning.com/boundless-psychology/chapter/trait-perspectives-on-personality/ (accessed March 30, 2022).

Hairer, M. (2021). *Amadeus Pro (version 2.8.8).* London, United Kingdom; HairerSoft.

Hanson, H. M. (1997). Glottal characteristics of female speakers: Acoustic correlates. *J. Acoust. Soc. Am.* 101, 466–481. doi: 10.1121/1.417991

Hart, S. D., and Hare, R. D. (1994). Psychopathy and the Big 5: correlations between observers' ratings of normal and pathological personality. *J. Pers. Disord.* 8, 32. doi: 10.1521/pedi.1994.8.1.32

Hillenbrand, J. (1987). A methodological study of perturbation and additive noise in synthetically generated voice signals. *J. Speech Lang. Hear. Res.* 30, 448–461. doi: 10.1044/jshr.3004.448

Hillenbrand, J., and Houde, R. A. (1996). Acoustic correlates of breathy vocal quality: Dysphonic voices and continuous speech. *J. Speech Lang. Hear. Res.* 39, 311–321. doi: 10.1044/jshr.3902.311

Hodges-Simeon, C. R., Gaulin, S. J., and Puts, D. A. (2010). Different vocal parameters predict perceptions of dominance and attractiveness. *Human Nat.* 21, 406–427. doi: 10.1007/s12110-010-9101-5 ncq

John, O. P., and Srivastava, S. (1999). The Big Five trait taxonomy: history, measurement, and theoretical perspectives. *Handbook Personal.* 2, 102–138.

Ladefoged, P. N., and Johnson, K. (2011). "Airstream mechanisms and phonation types," in *A course in phonetics.* Boston etc.: Wadsworth.

Laver, J. (1972). "Voice quality and indexical information readings," in *Communication in face to face interaction,* Laver, J., and Hutcheson, S. (eds.). London, United Kingdom: Penguin. p. 189–203.

Laver, J. (1980). The phonetic description of voice quality. *Cambridge Stud. Linguist. London.* 31, 1–186.

McCrae, R. R., and John, O. P. (1992). An introduction to the five-factor model and its applications. *J. Pers.* 60, 175–215. doi: 10.1111/j.1467-6494.1992.tb00970.x

Michalsky, J., and Niebuhr, O. (2019). Myth busted? *Challenging what we think we know about charismatic speech. AUC PHILOLOGICA* 2019, 27–56. doi: 10.14712/24646830.2019.17

Möbius, B. (2003). "Gestalt Psychology meets phonetics-An early experimental study of intrinsic F0 and intensity," in *Proceedings of the 15th International Congress of Phonetic Sciences, Vol. 3* (Barcelona: IPA and UAB), 2677–2680.

Murton, O., Hillman, R., and Mehta, D. (2020). Cepstral peak prominence values for clinical voice evaluation. *Am. J. Speech Lang. Pathol.* 29, 1596–1607. doi: 10.1044/2020_AJSLP-20-00001

Niebuhr, O., and Fischer, K. (2019). Do not hesitate!-unless you do it shortly or nasally: how the phonetics of filled pauses determine their subjective frequency and perceived speaker performance. *Interspeech.* 544–548. doi: 10.21437/Interspeech.2019-1194

Niebuhr, O., Skarnitzl, R., and Tylečková, L. (2018). The acoustic fingerprint of a charismatic voice-Initial evidence from correlations between long-term spectral features and listener ratings. *Proc. Speech Prosody.* 9, 359–363. doi: 10.21437/SpeechProsody.2018-73

Norman, W. T. (1963). Toward an adequate taxonomy of personality attributes: Replicated factor structure in peer nomination personality ratings. *J. Abnorm. Psychol.* 66, 574. doi: 10.1037/h0040291

Novák-Tót, E., Niebuhr, O., and Chen, A. (2017). A gender bias in the acoustic-melodic features of charismatic speech? *Interspeech.* 2248–2252. doi: 10.21437/Interspeech.2017-1349

Park, Y., Perkell, J. S., Matthies, M. L., and Stepp, C. E. (2019). Categorization in the perception of breathy voice quality and its relation to voice production in healthy speakers. *J. Speech Lang. Hear. Res.* 62, 3655–3666. doi: 10.1044/2019_JSLHR-S-19-0048

Parker, M. A., and Borrie, S. A. (2018). Judgments of intelligence and likability of young adult female speakers of American English: the influence of vocal fry and the surrounding acoustic-prosodic context. *J. Voice* 32, 538–545. doi: 10.1016/j.jvoice.2017.08.002

Pittam, J. (1985). *Voice quality: Its measurement and functional classification (Doctoral dissertation)*. Queensland University, Brisbane, QLD, Australia.

Pittam, J. (1987). The long-term spectral measurement of voice quality as a social and personality marker: a review. *Lang. Speech.* 30, 1–12. doi: 10.1177/002383098703000101

Pointer, N. F., van Mersbergen, M., and Nanjundeswaran, C. D. (2022). Listeners' attitudes towards young women with glottal fry. *J. Voice.* doi: 10.1016/j.jvoice.2022.09.007

Puts, D. A., Hodges, C. R., Cárdenas, R. A., and Gaulin, S. J. (2007). Men's voices as dominance signals: vocal fundamental and formant frequencies influence dominance attributions among men. *Evol. Hum. Behav.* 28, 340–344. doi: 10.1016/j.evolhumbehav.2007.05.002

Quené, H., Boomsma, G., and van Erning, R. (2016). Attractiveness of male speakers: effects of voice pitch and of speech tempo. *Proc. Speech Prosody.* 8, 1086–1089. doi: 10.21437/SpeechProsody.2016-223

R Core Team (2018). *R: A Language and Environment for Statistical Computing.* Vienna, Austria: R Foundation for Statistical Computing. Available online at: https://www.R-project.org/

R Studio Team (2018). *RStudio: Integrated Development for R.* Boston, MA: RStudio, Inc. Available online at: http://www.rstudio.com/ (accessed January 15, 2022).

Reetz, H., and Jongman, A. (2020). *Phonetics: Transcription, Production, Acoustics, and Perception.* New York: John Wiley and Sons.

Rosenberg, A., and Hirschberg, J. (2009). Charisma perception from text and speech. *Speech Commun.* 51, 640–655. doi: 10.1016/j.specom.2008.11.001

Scherer, K. R. (1979). Voice and speech correlates of perceived social influence in simulated juries. *Lang. Social Psychol.* 5, 88–120.

Shor, R. E. (1978). The production and judgment of smile magnitude. *J. Gen. Psychol.* 98, 79–96. doi: 10.1080/00221309.1978.9920859

Signorello, R., and Demolin, D. (2013). The physiological use of the charismatic voice in Political speech. *Interspeech*. 987–991. doi: 10.21437/Interspeech.2013-174

Tartter, V. C. (1980). Happy talk: Perceptual and acoustic effects of smiling on speech. *Percept. Psychophys.* 27, 24–27. doi: 10.3758/BF03199901

Tschinse, L. M., Asadi, A., Gutnyk, A., and Niebuhr, O. (2022). "Keep on smiling...? On the sex-specific connections between smiling duration and perceived speaker attributes in business pitches," in *Proceedings 11th International Conference of Speech Prosody* (Lisbon), 1–5. doi: 10.21437/SpeechProsody.2022-127

Tskhay, K. O., Zhu, R., Zou, C., and Rule, N. O. (2018). Charisma in everyday life: conceptualization and validation of the general charisma inventory. *J. Pers. Soc. Psychol.* 114(1), 131–152. doi: 10.1037/pspp0000159

Tull, R. G. (1999). *Acoustic Analysis of Cold-Speech: Implications For Speaker Recognition Technology and the Common Cold*. Evanston, Illinois: Northwestern University.

Vazire, S., Naumann, L. P., Rentfrow, P. J., and Gosling, S. D. (2009). Smiling reflects different emotions in men and women. *Behav. Brain Sci.* 32, 403–405. doi: 10.1017/S0140525X09991026

Vergauwe, J., Wille, B., Hofmans, J., and De Fruyt, F. (2017). Development of a Five-Factor Model charisma compound and its relations to career outcomes. *J. Vocat. Behav.* 99, 24–39. doi: 10.1016/j.jvb.2016.12.005

Warren, D. W., Hairfield, W. M., Seaton, D., Morr, K. E., and Smith, L. R. (1988). The relationship between nasal airway size and nasal-oral breathing. *Am. J. Orthod. Dentofacial.* 93, 289–293. doi: 10.1016/0889-5406(88)90158-8

Watterson, T., and Emanuel, F. (1981). Observed effects of velopharyngeal orifice size on vowel identification and vowel nasality. *Cleft Palate J.* 18, 271–278.

Weiss, B., and Moeller, S. (2011). "Perceptual dimensions of voice and way of speaking," in *Electronic Speech Signal Processing (ESSV)*. p. 261–268.

Wiener, H. J., and Chartrand, T. L. (2014). The effect of voice quality on ad efficacy. *Psychol. Market.* 31, 509–517. doi: 10.1002/mar.20712

Wolf, N. (2015). *Young Women, Give Up the Vocal Fry and Reclaim Your Strong Female Voice*. The Guardian. p. 24.

Wolk, L., Abdelli-Beruh, N. B., and Slavin, D. (2012). Habitual use of vocal fry in young adult female speakers. *J. Voice.* 26, 111–116. doi: 10.1016/j.jvoice.2011.04.007

Yuasa, I. P. (2010). Creaky voice: a new feminine voice quality for young urban-oriented upwardly mobile American women? *Am. Speech*. 85, 315–337. doi: 10.1215/00031283-2010-018

Ziegler, M., Danay, E., Schölmerich, F., and Bühner, M. (2010). Predicting academic success with the Big 5 rated from different points of view: Self-rated, other rated and faked. *Eur. J. Pers.* 24, 341–355. doi: 10.1002/per.753

frontiers | Frontiers in Communication

TYPE Original Research
PUBLISHED 23 March 2023
DOI 10.3389/fcomm.2023.972182

Check for updates

OPEN ACCESS

EDITED BY
Martti Vainio,
University of Helsinki, Finland

REVIEWED BY
Hemanta Kumar Palo,
Siksha O Anusandhan University, India
Dafydd Gibbon,
Bielefeld University, Germany

*CORRESPONDENCE
Oliver Niebuhr
✉ olni@sdu.dk
Ingo Siegert
✉ ingo.siegert@ovgu.de

SPECIALTY SECTION
This article was submitted to Language Sciences, a section of the journal Frontiers in Communication

RECEIVED 17 June 2022
ACCEPTED 24 February 2023
PUBLISHED 23 March 2023

CITATION
Niebuhr O and Siegert I (2023) A digital "flat affect"? Popular speech compression codecs and their effects on emotional prosody. *Front. Commun.* 8:972182. doi: 10.3389/fcomm.2023.972182

COPYRIGHT
© 2023 Niebuhr and Siegert. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

A digital "flat affect"? Popular speech compression codecs and their effects on emotional prosody

Oliver Niebuhr[1]* and Ingo Siegert[2]*

[1]Centre for Industrial Electronics, University of Southern Denmark, Sonderborg, Denmark, [2]Mobile Dialog Systems, Institute for Information Technology and Communication, Otto von Guericke University Magdeburg, Magdeburg, Germany

Introduction: Calls via video apps, mobile phones and similar digital channels are a rapidly growing form of speech communication. Such calls are not only—and perhaps less and less— about exchanging content, but about creating, maintaining, and expanding social and business networks. In the phonetic code of speech, these social and emotional signals are considerably shaped by (or encoded in) prosody. However, according to previous studies, it is precisely this prosody that is significantly distorted by modern compression codecs. As a result, the identification of emotions becomes blurred and can even be lost to the extent that opposing emotions like joy and anger or disgust and sadness are no longer differentiated on the recipients' side. The present study searches for the acoustic origins of these perceptual findings.

Method: A set of 108 sentences from the Berlin Database of Emotional Speech served as speech material in our study. The sentences were realized by professional actors (2m, 2f) with seven different emotions (neutral, fear, disgust, joy, boredom, anger, sadness) and acoustically analyzed in the original uncompressed (WAV) version and as well as in strongly compressed versions based on the four popular codecs AMR-WB, MP3, OPUS, and SPEEX. The analysis included 6 tonal (i.e. f0-related) and 7 non-tonal prosodic parameters (e.g., formants as well as acoustic-energy and spectral-slope estimates).

Results: Results show significant, codec-specific distortion effects on all 13 prosodic parameter measurements compared to the WAV reference condition. Means values of automatic measurement can, across sentences, deviate by up to 20% from the values of the WAV reference condition. Moreover, the effects go in opposite directions for tonal and non-tonal parameters. While tonal parameters are distorted by speech compression such that the acoustic differences between emotions are increased, compressing non-tonal parameters make the acoustic-prosodic profiles of emotions more similar to each other, particularly under MP3 and SPEEX compression.

Discussion: The term "flat affect" comes from the medical field and describes a person's inability to express or display emotions. So, does strong compression of emotional speech create a "digital flat affect"? The answer to this question is a conditional "yes". We provided clear evidence for a "digital flat affect". However, it seems less strongly pronounced in the present acoustic measurements than in previous perception data, and it manifests itself more strongly in non-tonal than in tonal parameters. We discuss the practical implications of our findings for the everyday use of digital communication devices and critically reflect on the generalizability of our findings, also with respect to their origins in the codecs' inner mechanics.

KEYWORDS

prosody, coded speech, emotional speech, phonetics, distant meetings, degraded speech

1. Introduction

1.1. What is flat affect?

The term "flat affect" comes from the medical field and describes a person's inability to express or display emotions in the same way as other people do, especially with regard to the intensity of the individual emotions, but also in terms of their differentiation. It already manifests itself at the level of the cognitive processes underlying emotional expressions. In principle, this makes the "flat affect" a condition in its own right, but in practice it often co-occurs with other conditions, such as autism or schizophrenia (Evensen et al., 2012).

A flat-affect condition concerns both verbal and nonverbal aspects of emotional communication. However, prosody is particularly impaired. For example, Gottheil et al. (1970) conducted an experiment in which participants rated either audio stimuli or only their transcribed texts from a flat-affect patient group and a control group. In the audio condition, both groups could be clearly differentiated by participants (with the control group stimulating significantly higher ratings of emotional intensity). By contrast, in the transcribed-text condition, i.e., without the phonetics and, in particular, the speech prosody, the participants were no longer able to differentiate between the two groups. In accord with that, Alpert et al. (2000) concluded from their study that it is not so much the use of emotional words (both positive and negative) that distinguishes flat-affect patients from a control group, but the (in)ability to express emotions nonverbally. Flat-affect patients, for example, cannot in the same way as a healthy control group convey emotions and their intensity levels by means of Fundamental Frequency (f0) inflection. In an earlier study, Alpert and Anderson (1977) have shown already that flat-affect patients also suffer from a reduced level and variability of voice intensity in their emotional statements.

The aim of the present study is to examine by means of a prosodic analysis whether popular speech compression codecs also make those acoustic patterns disappear partially or completely that allow listeners to perceive the emotional intensity of utterances (e.g., such that they become to some degree acoustically indistinguishable from neutral matter-of-fact utterances) and/or the emotional categories of utterances (e.g., such that the acoustic profiles of different emotions merge to some degree). We call this a "digital flat affect"; digital because the flat affect does not come from the speaker him/herself, but is caused by the codec as an artifact in the digital communication channel. The background to this question is explained below.

1.2. What is charisma?

Phonetic research on speaker impact or perceived speaker charisma began with the seminal papers of Touati (1993), Strangert and Gustafson (2008), and Rosenberg and Hirschberg (2009). The latter have defined charisma as "the ability to attract and retain followers without the benefit of formal authority" (p.640). We do not dispute the truth in this definition, but we agree with Antonakis et al. (2016) that charisma should not be defined in terms of its outcome, i.e., its effects. Such definitions easily become circular. They also ignore a key insight of charisma research over the past few decades, namely that charisma is not something a person has, but something a person does. It is a learnable signaling system, see also Antonakis et al. (2011) on teaching charisma. Antonakis et al. (2016) therefore, define charisma as "values-based, emotion-laden leader signaling" (p. 304).

We follow this definition. It focuses on our research subject: signals; and it leaves room for the same signals to be rated sometimes more and sometimes less charismatic (e.g., by different listeners or in different contexts) because it does not define charisma from an effects' perspective. The addition of "values-based" in the definition of Antonakis et al. expresses that, for charisma to unfold its intended effects, the same (e.g., moral) values must be shared by speaker and audience or, in more general terms, by sender and recipient (cf. Kelman, 1961). In telemarketing contexts like those investigated by Chebat et al. (2007), for instance, such shared values are often established by the call-center agent, who starts the dialogue with a rhetorical question like "Don't we all want to save taxes?". Finally, and this is the crucial point here, the definition of Antonakis et al. (2016) links charisma to the expression of emotions (see Section 1.4).

In addition to the obvious verbal strategies such as three-part lists and metaphors (cf. the "Charismatic Leadership Tactics" of Antonakis et al., 2011), the complex bundle of charismatic signals also includes facial expressions and gestures (Keating, 2011), speech prosody (Rosenberg and Hirschberg, 2009), choice of words (Tur et al., 2021), clothing (Furnham et al., 2013; Karabell, 2016) as well as gender, age, and height (Grabo et al., 2017; Jokisch et al., 2018; Maran et al., 2021). The present study deals with the phonetic aspects of charisma, more precisely with the effects of charismatic prosody. Section 1.3 provides a research overview on this area.

1.3. Previous research on prosodic charisma and speech compression

Siegert and Niebuhr (2021a) carried out a perception experiment in which listeners rated short stimulus utterances that were presented both as uncompressed originals and as variants heavily reduced by four popular speech-compression codecs: AMR-WB, MP3, OPUS, and SPEEX. The rating was performed against the background of perceived speaker charisma, i.e., by using 10 scales that are associated with charisma according to a meta-analysis of previous studies. The results of Siegert and Niebuhr showed that, compared to the uncompressed baseline, speech compression had a significantly negative impact on the speakers' ratings along the charisma traits— and that, moreover, this negative effect was significantly more pronounced for female speakers than for male speakers. More pronounced means that the negative effect involved more codecs (including OPUS, which even improved the charisma ratings of male speakers) as well as more rating scales, and it additionally lowered the female speakers' ratings more strongly on many scales. For example, under speech compression, women lost more of their perceived trustworthiness, persuasiveness, and likability than men.

Later, Siegert and Niebuhr (2021b) showed in a supplementary investigation of the compressed and original stimulus material of Siegert and Niebuhr (2021a) that an automatic prosodic analysis yields significantly different results for the compressed and original stimuli. These differences concerned f0, intensity, and exponents of voice quality, but not durational measures. Crucially, if the different automatic measurement results were interpreted in view of known correlations between prosody and perceived speaker charisma, then the codecs' acoustic differences matched well with the codecs' perceptual effects in Siegert and Niebuhr (2021a). In other words, the codec-induced reduction of perceived speaker charisma in Siegert and Niebuhr (2021a) is consistent with the codec-induced acoustic-prosodic measurement changes in Siegert and Niebuhr (2021b).

This finding has practical everyday implications beyond a mere scientific relevance, because acoustic charisma can influence the opinions and behavior of recipients and, thereby, a speaker's professional success. For example, Antonakis et al. (2016) report that the enhanced charisma signals of charismatically trained speakers can lead to 17% more output among workers and increase people's contributions to the common good by up to 19%. Fischer et al. (2019) showed further that such effects can be transferred to machines, solely based on enhancing their prosodic cues to charisma: If one transfers the tone-of-voice patterns of Steve Jobs, Apple's former exceptionally charismatic CEO, to a robot, leaving all other features of the speech synthesis constant, then this robot becomes significantly more successful in making human interaction partners fill out longer questionnaires, eat healthier food, and take detours in traffic (Niebuhr and Michalsky, 2019); and when such a robot is used as a moderator in creativity workshops, then these workshops end with significantly more and better ideas, and give participants a more positive mindset in this idea-generation process (Fischer et al., 2022). Gregory and Gallagher (2002) were able to correlate all outcomes of the presidential elections in the USA between 1960 and 2000 with a single prosodic measure derived from f0 and the lower spectral harmonics of the candidates' voices. Niebuhr (2021) found correlations between a prosody-based charisma score and the oral exam grades of Danish university students. Pentland (2008) was able to predict with 87% accuracy only from nonverbal signals who would win a business-plan competition; a finding that stimulated the newspaper headline "no charisma—no funding" in a silicon-valley magazine. Similarly, but only by means of five tonal and five non-tonal prosodic parameters, Niehof and Niebuhr (2022) were recently able to determine 70–80% correctly which investor pitches win or lose in major German pitching contests.

So, the prosody of the sender (be it a human or a robot) is one of the key components in creating charismatic effects. Further supporting evidence for this statement comes from the field of speech technology. It has been repeatedly shown that machine-learning models benefit most from prosody (rather than from video or text information) when trying to replicate charisma-related human behavior; see the studies by Chen et al. (2014), Wörtwein et al. (2015), and Kimani et al. (2020) on modeling ratings of public-speaking performances, or the study by Amari et al. (2021) on modeling the success of persuasion strategies in consultant-client dialogues.

1.4. Charisma and the expression of emotions

It was House (1977) who, in the first psychological theory on charismatic leadership, made the expression of emotions a core element of charismatic impact. Also, according to Bass (1990), charismatic speakers are, amongst other things, "determined, self-confident, and emotionally expressive" (p.220), see Antonakis et al. (2016) for a historical summary. Emotions are important because they create the motivational basis for the transfer of goals, ideas, or instructions from the charismatic sender to his/her recipients *via* the emotional-contagion effect, see Bono and Ilies (2006). In the words of Sy et al. (2018), a charismatic impact relies heavily on "eliciting and channeling follower emotions" (p. 58).

In fact, there are many studies showing the positive influence of emotions on the perception of a person as charismatic—as well as the importance of displaying emotions for business success, see the overview in Humphrey et al. (2008). Kisfalvi and Pitcher (2003) provide case studies to demonstrate the connection between the emotional nature of a CEO and the economic success of his/her company and team. Similarly, significant correlations can be established between crowdfunding success on the one hand and the emotional intensity of the speech and non-speech content in the corresponding crowdfunding videos on the other (Kim et al., 2016; Liu et al., 2018); see also the related study of Zhao et al. (2022) who showed that it is the linguistic rather than the visual emotional content that determines the success of a campaign. Davis et al. (2017) also identified the emotional passion (of entrepreneurs) as a key factor for campaign success and, in this, particularly stressed the role of non-verbal features.

Passion as an element of perceived speaker charisma is often intuitively associated with positive emotions (Damen et al., 2008). There is, however, accumulating evidence that this intuitive idea is too simplistic. D'Errico and Poggi (2022) recently showed for a sample of Italian politicians that the expression of negative emotions can be interpreted as a sign of power and strength which, in challenging times, can create a charismatic effect as well (cf. also Gooty et al. 2010). D'Errico et al. (2019) already emphasized for a perception experiment with a set of well-known politicians that "activating emotions, joy and anger, increased the positive evaluation of the politician, whom participants perceived as more [...] charismatic" (p. 671). Zhao et al. (2022) pointed out in their research overview that funding (e.g., donation) campaigns can be successful based on negative emotions. Also, the results of Allison et al. (2022) suggest that it is primarily the high arousal of an emotion that makes speakers more charismatic, not the emotion's specific valence.

As with charisma, prosody plays an important role in expressing emotions. Different emotions have different prosodic patterns, see the detailed studies by Mozziconacci (2001), Gobl and Chasaide (2003), and Bänziger and Scherer (2005)— and the prosodic parameters within each pattern are unidirectionally related to the intensity of the emotion (e.g., Audibert et al. 2010 and see Rosenberg and Hirschberg 2009 for charisma). How central prosody is for the expression of emotions can be seen from two sets of robust findings. Firstly, based on prosodic factors alone, speech-recognition systems are able to identify the emotional category in a

speech utterance with sometimes more than 90% accuracy (Luengo et al., 2005; Kao and Lee, 2006; Wang et al., 2008), see also the overview in Koolagudi and Rao (2012). Secondly, studies show that listeners are still able to reliably identify emotion categories in foreign-language stimuli, i.e., in the absence of any lexical cues (and visual cues), albeit worse than for their own native language and worse for the valence than for the arousal dimension of emotions, see Feraru et al. (2015) for a recent overview.

1.5. Connecting the threads: Toward the assumption of a digital flat affect

Prosodic patterns are an important source of perceived speaker charisma; and the compression of speech signals by codecs can weaken these prosodic patterns, both on the perceptual side and on the acoustic side—the latter in a way that matches with the former perceptual effects. The expression of charisma is closely related to the expression of emotions; and the perception of emotional categories and their intensity is also based in large part on prosodic patterns. It would therefore be reasonable to assume that emotional prosodic patterns are also negatively affected by codecs and that this results in what was referred to in 1.1 as a digital flat affect (note that studies on the connection between speech compression and emotion recognition are numerous, but so far only related to automatic emotion recognition and/or, if conducted with human listeners, then aimed at testing compression devices, such as for hearing aids, see Goy et al. 2016; Reddy and Vijayarajan 2020).

Taking up this flat-affect assumption, Niebuhr and Siegert (2022) extended their investigations from charisma to emotions. As before (see Section 1.3), the first step was a perception experiment. The experiment used the same charisma scales as in Siegert and Niebuhr (2021a), but now with a set of enacted emotional speech stimuli from the Berlin Database of Emotional Speech (EMODB) (Burkhardt et al., 2005). The stimuli represented the seven categories fear, disgust, boredom, neutral, anger, joy, and sadness. As in Siegert and Niebuhr (2021a), all emotional stimuli were compressed to various degrees by the codecs AMR-WB, MP3, OPUS, and SPEEX and then rated by listeners.

On this basis, Niebuhr and Siegert (2022) tested the ability of listeners to differentiate emotions in codec-compressed speech stimuli. Furthermore, Niebuhr and Siegert tested whether positive emotions generally trigger higher and negative emotions lower charisma rating levels or whether, in line with the above findings, the rating levels related to charisma tend to be independent of emotional valence. The study comes to three main conclusions: First, the applied charisma-related rating scales are generally able to significantly differentiate the seven emotion categories tested. Second, in line with the findings summarized above, valence was not the sole determining factor for rating levels. Negative emotional stimuli were thus also able to evoke high rating levels along the charisma-associated scales. Moreover, it was not the emotion with the lowest arousal, boredom, that triggered the lowest charisma ratings. Rather, the most uncharismatic emotion was disgust, not least because it evoked low charisma ratings regardless of speaker sex, while, e.g., fear made men sound more charismatic than women, and sadness made women sound more charismatic than men.

Third, both the type of audio codec and the strength of signal compression significantly impaired emotion differentiation along the rating scales. The worse the audio quality, the more the emotional impressions merged along the scales, even for popular codecs like MP3. For example, when being strongly compressed, high-arousal emotions such as joy became, in terms of the listener ratings, indistinguishable from low-arousal emotions such as boredom and sadness. Niebuhr and Siegert (2022) have thus provided the first perceptual evidence of a codec-induced digital flat affect. Analogous to the study by Siegert and Niebuhr (2021b), the present study is intended to shed light on the acoustic-prosodic side of this perceptual evidence.

1.6. Codecs and their influence on prosody

Speech compression is heavily used in modern mobile systems. It reduces the transmission bandwidth while maintaining speech intelligibility (Albahri et al., 2016). This allows engineers to reduce transmission delay as well as memory and storage capacities. Two different compression approaches have been developed. Psychoacoustic modeling is mainly used for transparent music compression, aiming to simultaneously reduce the file size and preserve all audible acoustic information. Prominent codecs are MP3, Ogg Vorbis, or Advanced Audio Coding (AAC). For real-time speech conversation applications, the Analysis-by-Synthesis approach predominates, designed to obtain an acceptable intelligibility at limited bandwidths, under real-time conditions and, if required, even at very low bit-rates. Important codecs using Analysis-by-Synthesis are SPEEX, AMR-(NB/WB/WB+), GSM-FR, or EVS. Furthermore, some codecs can switch between speech and audio compression modes based on signal content. Several codecs have been developed coordinating both approaches and offering various computational and intelligibility improvements (Vary and Martin, 2006). The Opus codec offers a hybrid mode in order to further improve speech intelligibility at low bit-rates.

Table 1 depicts an overview of parameters for selected audio codecs. Most Analysis-by-Synthesis codecs strive for perceptually optimizing the decoded (synthesis) signal in a closed loop. In order to achieve real-time processing, the algorithms involve three processing steps. This method is known as CELP: First, line spectral pairs (LSPs) are computed and quantized. The remaining properties of the speech signal are then estimated by using a codebook of fixed (i.e., hard-coded) quantization vectors, with a typical size of 16 bits. SPEEX uses CELP and AMR-WB as well, albeit an optimized version (i.e., an optimized codebook) referred to as ACELP for which frame rates are fixed at 20 ms. The codebook can contain more than 50 bits of entries. Another variant of this approach is SILK, used in the Opus codec. It also uses Analysis-by-Synthesis approaches but with variable frame sizes (10, 20, 40 and 60 ms), and it encodes up to 8 kHz. For full-bandwidth audio signals, Opus implements CELT, a high-quality, low-delay audio codec. CELT is based on the MDCT and supports frame sizes of 2.5, 5, 10, and 20 ms as well as wideband (8 kHz), super-wideband

TABLE 1 Overview of important parameters of selected audio codes.

	Frame	Frame size [ms]	Speech coding	Data size	Look-ahead [ms]
SPEEX	Fixed	30/34 (8/16 kHz)	CELP	16 bits	5
AMR-WB	Fixed,	20	ACELP	up to 50 bits	5
OPUS SILK	Variable	10, 20, 40 and 60	SILK	Unknown	5
OPUS CELT	Variable	2.5, 5, 10 and 20	CELT	Unknown	2.5
MP3	Fixed	24	MDCT	576 coefficients (192 for transients)	528 samples

(12 kHz) and full-band (20 kHz) audio processing. Furthermore, the hybrid mode offered by OPUS uses SILK for frequencies up to 8 kHz and CELT to encode the frequencies above 8 kHz, but only based on frame sizes of 10 and 20 ms.

By contrast, the psychoacoustic modeling employed in MP3 compression strives to discard exactly those parts of the original sound signal that are considered beyond human auditory resolution. To that end, data frames of 24 ms are quantified using a polyphase filter bank with 32 equal frequency bands (less precisely perceptible frequency ranges with lower assumed resolution). Using equally wide frequency bands simplifies the filters but does not reflect human hearing, for which the resolution is non-linearly associated with frequencies. Afterwards, MDCT is applied, resulting either in 576 spectral coefficients (long blocks, 32 taps x 18 sub-bands) or three times 192 spectral coefficients (short blocks, 32 taps x 6 sub-bands). The 192-coefficient approach is applied—if there is a transient[1]—to limit the temporal spread of quantization noise accompanying this transient. Then, a quantization is used in which adjacent frequency bands are combined into groups of 4 to 18 bins and share a scaling factor determining the accuracy of the (in this case non-linear) coding for these frequency bands. Finally, the scale factors and the quantized amplitudes of the individual frequencies are Huffman-coded using fixed code tables (Lutzky et al., 2004). The relatively small block window decreases the coding efficiency, and the short time resolution can be too low for highly transient signals. That is, it can cause smearing of percussive sounds. Furthermore, the tree structure of the filter bank increases pre-echo problems and the combination of two adjunct filter banks creates aliasing problems, that are only partially compensated for at a later stage of the encoder.

Although a number of studies investigated the general impact of codec compression on spectral quality and acoustic features (Byrne and Foulkes, 2004; Guillemin and Watson, 2009; Siegert et al., 2016), the effects on the preservation of emotions, nonverbally conveyed ones in particular, have rarely been addressed (Albahri et al., 2016; Jokisch et al., 2016). Especially, the preservation of nonverbal emotional cues under low bandwidths is under-researched. In the following, some related findings for the codecs used in the current investigation are summarized.

Pollak and Behunek (2011) studied the masking and attenuation of frequency components under MP3 compression and found that this can lead to the suppression of a sound segment, and sometimes to inter-word pause shortening, which can reduce the perceived fluency of speech utterances. Furthermore, the masking and attenuation of frequency components can worsen the estimation error of power spectra in the output of the filter bank. This is especially crucial for MFCC and PLP features (Pollak and Behunek, 2011). The authors performed their experiments with quite high MP3 bit-rates (160 kbit/s). Thus, it can be assumed that the negative effects they reported would manifest themselves still more clearly at very low bit-rates. In Bollepalli (2013), a range of MP3 bit-rates from 8 to 160 kbit/s is investigated. The aim was to investigate the influence of compression on acoustic-prosodic features important for speech synthesis, such as f0, harmonics-to-noise ratio (HNR), line spectral frequencies (LSF) of voice source and vocal tract, harmonic amplitude difference (H1-H2), and normalized amplitude quotient (NAQ). For each compression rate, the relative error was determined between the parameter value computed from the uncompressed and compressed sound. Bollepalli (2013) observed that, for f0, the error was negligible at high bit-rates (64 kbit/s or more) and less than 5% on average at low bit-rates. However, it is noteworthy that lower bit rates not only made the magnitude of the error increase, but also its range. That is, in connection with other (spectral) signal properties, f0 errors are pushed in specific directions. For HNR, the error was considerably larger, particularly in high frequency bands, thus suggesting a better preservation of the harmonic structure in low-frequency bands. For LSF, an inverse observation was made. Both voice source and vocal tract measures showed the strongest increases in errors at low frequencies. Regarding how H1-H2 and NAQ differ as a function of original and compressed signals, Bollepalli et al. showed that the goodness of correlation between original and compression signal declined gradually with decreasing bit-rates. Especially H1-H2 suffers greatly from high compression, while NAQ remains relatively stable.

Another study of van Son (2005) analyzed the influence of codec compression on pitch, formant levels, and the spectral center of gravity (CoG). It was concluded that compression added a Root-Mean-Square (RMS) error of less than 1 semitone to vowel mid-point pitch, formant, and CoG measurements in general. Only for the lowest tested bit-rate (40 kbit/s), rather large CoG errors emerged in sonorous consonants and fricatives ($>$ 2 semitones).

In summary, we see that codec compression has an effect on (automatic) prosodic measurements. The effect includes both f0 and spectral properties. However, f0 properties seem to be

1 A transient is a high amplitude, short-duration sound at the beginning of a waveform.

less affected than spectral properties, and among the latter, high-frequency spectral information seems to suffer more from compression than low-frequency spectral information, albeit both are affected. Furthermore, measures that integrate several signal features such as CoG are more strongly affected, particularly if these features are narrowly defined local characteristics as in the case of H1-H2. It is reasonable to assume that, when H1-H2 is particularly affected, that the same also applies to f0 range measures, which also represent a difference value between two local characteristics. For f0 measures, especially those related to levels, we can additionally expect an interaction with spectral characteristics, i.e., with the voice quality and, thus, with the type of emotion. That is, depending on the type of emotion, it is possible that the f0 measures get pushed in difference directions by compression. Thus, the question is not if codec compression affects the prosodic profiles of emotions, but only to what degree for the individual codecs. Based on the starting point for the present study outlined in Section 1.5, we therefore pursue three research questions, which we present in Section 1.7—and which we address according to Section 1.5 and Section 1.6 with reference to both tonal (f0) and non-tonal parameters.

1.7. Questions

The present exploratory study aims to see behind the curtain of the perception results of Niebuhr and Siegert (2022) by means of a multi-parametric acoustic-prosodic analysis. This lays the foundation for three fields of knowledge, which are represented by the following three questions:

- How does speech signal compression, which is ubiquitous in the rapidly growing use of digital communication, affect the acoustics of different emotions?
- How is emotion acoustics related to listener ratings of perceived speaker charisma and related traits?
- Following up on question (2), how is the merging of emotions along the charisma scales reflected in prosodic parameters?

Note that all three questions have a gender-specific dimension that is taken into account here. Note further that the many emotions studied, and their compressed manifestations, provide an acoustic range that will also help us in subsequent studies to better understand which acoustic parameters and parameter classes are related to perceived charisma and how.

2. Study design

2.1. Stimuli

For the stimulus material, we relied on the Berlin Database of Emotional Speech, EMODB (Burkhardt et al., 2005). EMODB is a benchmark dataset of high-quality recordings of expressive emotions, similar to those used for analyzes of, e.g., speech synthesis (Steidl et al., 2012), spectral and temporal changes in emotional speech (Kienast and Sendlmeier, 2000), influences of different room acoustics (Höbel-Müller et al., 2019), or automatic emotion recognition (Schuller et al., 2009). EMODB is for German and has the advantage that it is established in the area of speech technology (Ververidis and Kotropoulos, 2006), with the largely unanimous identification of the emotional categories in the stimuli being repeatedly proven (Schuller et al., 2011; Weiss et al., 2013; Bhangale and Mohanaprasad, 2021). The emotions are: anger (Ärger), fear (Angst), joy (Freude), sadness (Trauer), disgust (Ekel), boredom (Langeweile), and neutral, a reference condition.

Studies on emotional expressions have in the past also relied on natural, real-life emotions, e.g., Tarasov and Delany (2011). Enacted emotions as in the EMODB are sometimes criticized as being exaggerated or prototypical and, thus, less ecologically valid. We agree that there are nowadays smart and indirect methods to elicit a range of emotions from native speakers (e.g., Pfitzinger and Kaernbach, 2008; Pfitzinger et al., 2011; Niebuhr et al., 2015). Yet, we would counter that, firstly, the underlying assumption that authentic everyday emotions always convey the speaker's actual feelings is unverifiable and cannot be taken for granted. In the case of negative emotions that are strongly regulated by society, this is actually rather unlikely. In this respect, authentic emotions are also not unproblematic in terms of their ecological validity. They probably mostly show a positivity bias in both their level of expressivity as well as in the frequency of occurrence of the displayed emotional categories. Secondly, a comparative study by Scherer (2013) shows that, at least for two basic types of emotions like happy and sad, the prosodic profiles of authentically stimulated and professionally enacted emotions came out largely congruent. Thirdly, it is essential for the present study that the emotions are realized on similar, if not identical, verbal statements. Such a high degree of experimental control is unattainable with naturally occurring, authentic emotions.

The crucial point, however, is that this authenticity was not necessarily desirable for the purposes of the present study. Rather, for our question of whether speech-compression codecs alienate or reduce the acoustic fingerprints of emotional speaker expressions, it makes sense to start from clearly pronounced, perhaps even exaggeratedly produced, emotional prototypes—and to include strongly contrastive emotions like joy, sadness, and disgust, the latter two of which are hard to find in authentic field data. On this basis, the chance to detect codec effects increases; and if there are already serious limitations in the acoustic discrimination of emotions for such clear speech material, then these discrimination limitations should be all the more problematic for blurrily separated, authentic emotions uttered outside of studio recording conditions (at least on an acoustic level, i.e., detached from the string of words and the conversational context). Based on EMODB, we can examine codec effects independently of effects of recording equipment and gain, speaker, and acoustic conditions.

Each utterance was contextualized for the speaker in the recording session, e.g., happiness after winning a large amount of money in the lottery or sadness caused by losing a very good friend or relative. The speakers were asked to put themselves in the given situation and then project that emotion onto the respective utterance. Afterwards, by conducting a perception test with 20 listeners, the final 494 single-sentence utterances have achieved naturalness ratings higher than 60%, and an emotion recognizability over 80% (Burkhardt et al., 2005).

For our study, we selected the same subset of 2 male and 2 female speakers (#11, #15 and #13, #14 in the corpus respectively) as in our previous studies (Siegert and Niebuhr, 2021a,b; Niebuhr and Siegert, 2022), but this time we utilized the neutral versions as well as the emotional version of the uttered sentence. This resulted in a total of 5 (codecs) *26 (samples) = 130 samples. For each emotion, we have 1 sample per speaker per emotion. Besides the obvious advantages of using the same set of speakers and utterances as in previous studies (the latter were focused on the emotionally neutral utterance renditions, though), four speakers sufficed for the purposes of the present study.

The small sample size may raise skepticism as to the generalizability of our findings. However, note that our study was not a classic speech production study in which, for example, meaningful patterns were to be derived from a speech acoustics that varies according to inter-individual anatomies and physiologies as well as due to trade-offs of redundant acoustics cues in the speech code. Such studies need large speaker samples to be able to separate the meaningful patterns from the performance "noise" in which they are embedded. In contrast, in our study, we apply a constant processing procedure in the form of each speech-compression codec to speech signals. This application always leads to the same result for the same speech signal, whereas, for example, a repetition of the same utterance, even by the same speaker, never leads to the same result. In this sense, the EMODB utterances are in our study design only demonstration objects for the acoustic consequences of applied codecs. In principle, the emotional utterances of a single speaker would have sufficed for this purpose. However, we wanted to take the factor speaker sex into account and, in addition, included two speakers per sex for a sufficient number of items in the statistical analysis (apart from the fact that we also wanted to use the same speakers and utterances as in Siegert and Niebuhr 2021a,b; Niebuhr and Siegert 2022). Of course, with a sample of 2x2 speakers, we cannot represent the full range of inter-individual and between-sex variations in the expression of emotional utterances; and although this is a very relevant follow-up question, it is not the aim of the present study to analyze the effects of speech-compression codecs on such phonetic variations within individual emotional categories. In this first step, our basic goal is to analyze the effects of speech-compression codecs on the acoustic differences between the emotional categories themselves. Not least because of this, it was important that we started from clear, prototypically produced emotions and included strongly contrastive emotions like fear, sadness, disgust, and joy.

2.2. Utilized audio codecs

The main purpose of applying speech compression for mobile communication is to reduce the bandwidth for transmission, the transmission delay as well as the required system memory and storage (Maruschke et al., 2016; Siegert et al., 2016). Several codecs have been developed to meet various applications with different quality requirements, aiming to retain the speech intelligibility (ITU-T, 1996, 2014; Maruschke et al., 2016). To be in line with the previous investigations by the Siegert and Niebuhr (2021a,b), and Niebuhr and Siegert (2022), the same four codecs were utilized here.

Adaptive Multi-Rate Wideband (AMR-WB) is a high-quality audio compression format mainly used in mobile communications (ITU-T, 2003). Due to the processing of a wider speech bandwidth (50–6,400/7,000 Hz), this codec is also known as "HD Voice" and Voice over LTE (VoLTE) as AMR-WB extended the previously usual telephone quality of 3.4 kHz. We chose a bit-rate of 12.65 kbit/s, which is intended for pure speech signals (ITU-T, 2003). For this bit-rate, all signals above 6,400 Hz are neglected and replaced by sounds of the lower frequency data (75–6,400 Hz) along with random noise (in order to simulate the high frequency band).

MPEG-1/MPEG-2 Audio Layer III (MP3) is a popular lossy full band audio codec (Brandenburg, 1999). It uses perceptual coding for audio compression: certain parts of the original signal, considered to be beyond auditory resolution, are discarded. Besides its usage for music storage, lower bit-rates (16 kbit/s) are used to encode audio dramas (Ahern, 2020).

OPUS is an open-source lossy audio codec usable for both speech and music (Valin et al., 2012). It is nowadays used in many communication tools, as Zoom for video conferencing or WhatsApp for voice messages (Hazra and Mateti, 2017). OPUS further offers a hybrid mode to improve the speech intelligibility at low bit-rates by enriching the synthesized signal with characteristics represented by a psychoacoustic model (Valin et al., 2013). The application of the hybrid mode can be controlled by the bit-rate, which was 34 kbit/s in our case.

SPEEX is an open-source full band speech codec for Internet applications requiring particularly low bit-rates (Xiph.Org Foundation, 2014). It is also used as a speech codec in common voice assistant platforms (Caviglione, 2015). The encoding is controlled by a quality parameter that ranges from 0 (worst) to 10 (best). In our study, we used 0 (i.e., 3.95 kbit/s).

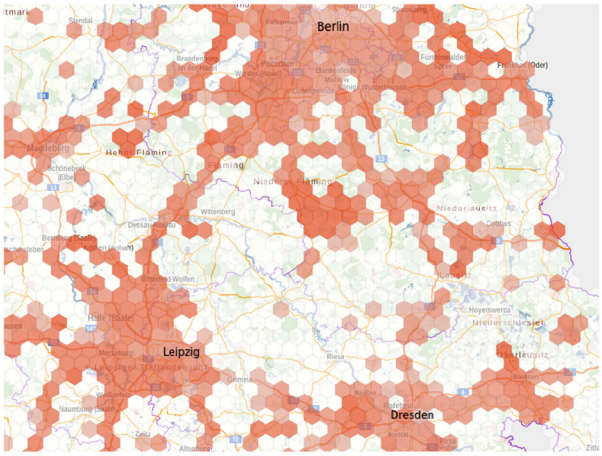


FIGURE 1
Mobile network coverage (5G) for the Vodafone Carrier Network as of June 2022. Map taken from Bundesnetzagentur (2022).

Another reason for the limitation to low bandwidths is the fact that sufficient data transmission bandwidth cannot be expected everywhere and in every situation, or service providers limit the bandwidth. To illustrate the first case, Figure 1 shows the "Map of mobile service availability" from the German Federal Network Agency. It illustrates very well the network coverage in German metropolitan areas and the lack of coverage in the rural regions. Figure 2 also shows the map of cellular dead zones. Here, it is noticeable that such cellular dead zones occur in major cities as well as in rural areas. In the large cities, the cause is usually infrastructure congestion due to too many users logged on at the same time, while in rural areas it is more likely to be due to a lack of suitable infrastructure. This justifies the usage of low bitrates for the current investigation and to analyze how well the selected codecs perform under these (imperfect) circumstances.

All 26 uncompressed stimuli have been compressed employing each of the four presented codecs at the specified bit-rate (AMR-WB: 12.65 kbit/s, MP3: 16 kbit/s, OPUS: 34 kbit/s, SPEEX: 3.95 kbit/s). This resulted in 104 compressed stimuli. The total number of stimuli in our experiment was hence 26 + 104 = 130 stimuli.

2.3. Procedure

The uncompressed WAV originals and the audio files derived from them through compression codecs were subjected to an acoustic-prosodic analysis. In phonetic research, prosody represents "an umbrella term used to cover a variety of interconnected and interacting phenomena, namely stress, rhythm, phrasing, and intonation." This definition by Arvaniti (2020) is functionally oriented, and although it excludes paralinguistic forms and functions like emotions and attitudes, it already suggests that prosody is far more than the fundamental frequency (i.e. f0) alone. It is true that f0, as the main acoustic correlate of perceived pitch, is perhaps most frequently examined in prosody research, not least because it is supposedly easy to measure (cf. Niebuhr et al. 2020). Nevertheless, prosody includes three further phenomenological dimensions. In phrasing, for example, duration plays a major role (with regard to intonation movements or phrase-final segmental units such as the syllable, e.g., Chavarría et al. 2004); and prominence as the basic building block of stress and rhythm is also encoded *via* changes in intensity and voice quality (Terken and Hermes 2000; Ludusan et al. 2021). Campbell and Mokhtari (2003) argue on empirical grounds "that voice-quality should be considered as the 4th prosodic parameter, along with pitch, power, and duration" (p.2417). In our study, we follow this phonetic (and hence pre-phonological and pre-functional) understanding of prosody of Campbell and Mokhtari. That is, by prosody we mean a bundle of phenomena constituted of four dimensions, namely f0, duration, intensity (or power) and voice quality—or their primary perceptual correlates pitch, length, loudness, and timbre.

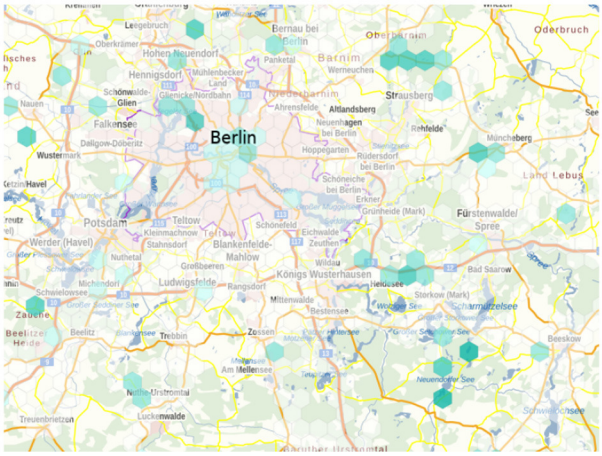


FIGURE 2
Map of cellular dead zones for Vodafone Carrier Network. Reported between Apr and Jun 2022. Map taken from Bundesnetzagentur (2022).

Our acoustic analysis covers all of these prosodic dimensions through multiple measurement parameters, with the exception of duration. This is because the previous study by Siegert and Niebuhr (2021b), which was concerned with the neutral statements of EMODB, could not find any compression effects on durational features of the signal. This makes sense, given that it is essential for every codec (e.g., in video call conditions) to leave the time axis of the voice signal untouched. Accordingly, in the present acoustic analysis we have also omitted all parameters that reflect a change in prosodic dimensions over time—not because we are not convinced of the functional relevance of timing and slopes. On the contrary, the first author intensively studied the functional relevance of variation in slope (Barnes et al., 2012) or contour shape in general (Landgraf, 2014). However, there are neither empirical nor conceptual indicators that would motivate measuring these parameters in our study. On the other hand, there is compelling evidence showing that, based on holistic measures of descriptive statistics, such as means, ranges, and standard deviations, emotional profiles can be properly characterized and distinguished—to the extent that statistical models like the Linear Discriminant Analyzes (LDAs) perform at eye level with human listeners in terms of the classification accuracy of emotions. Lausen and Hammerschmidt (2020) provide a recent summary of the relevant works.

The f0-related tonal dimension of prosody was represented in the acoustic analysis by six parameters: (i) mean f0, (ii) absolute, full f0 range (span), (iii) narrow f0 range that disregards the upper and lower 10% of the measurements (p90 range), (iv) f0 variability (standard deviation, f0 std), (v) minimum f0, and (vi) maximum f0.

For the non-tonal dimensions of prosody, the acoustic analysis took seven parameters into account: (i) mean intensity (RMS. dB), (ii) speaking rate (syll/s) as well as (iii) mean H1-H2 and (iv) mean H1-A3 (both f0-corrected), and (v-vii) the mean levels of the first three formants (F1, F2, F3). Note that speaking rate may at first glance appear as a duration parameter, and to some degree it is of course one. But, the rate is estimated in terms of the rising-falling intensity peaks that are assumed to underlie each syllable, and it is the variability and separability of these intensity peaks that we are interested in here. The two amplitude-difference

measures H1-H2 and H1-A3 are typical voice-quality parameters. H1-H2 is the difference between the first and the second harmonic amplitude of the spectrum. H1-A3 is the amplitude difference of the first harmonic and the harmonic closest to the third formant frequency. Both measures intend to estimate the speaker's spectral tilt. Menezes and Maekawa (2006) argue that H1-H2 reflects the open-quotient of the vocal-fold vibration cycle, whereas H1-A3 is associated with the abruptness of glottal closure. Therefore, H1-A3 may be more sensitive in capturing loudness- or effort-induced changes in voice quality, whereas H1-H2 performs well at representing differences between modal and breathy voices. Both measures have been found to be relevant for characterizing the acoustic profiles of emotions (Banse and Scherer, 1996; Liu and Xu, 2014).

Furthermore, mean F1, F2, and F3 seem to be unrelated to prosody. In fact, when measuring local formant-frequency targets or formant ranges, then the formant frequencies refer to segmental rather than prosodic characteristics, see, for example, the explanations in Ménard et al. (2007). However, when being measured in a literally suprasegmental way as mean resonance-frequency levels across entire utterances, the formant frequencies F1-F3 (and F4 too, see Rendall et al. 2005) become measures of voice quality. More specifically, they represent characteristics of vocal-tract length (or perceived speaker size), of the hollowness or brilliance of the voice, and of the speaker's articulatory setting, i.e., "vocal tract configurations that the articulators—including the jaw, larynx, pharynx, tongue, and lips—tend to be deployed from (and return to) in speech production" (Pratt, 2020:331). Such settings give all consonants and vowels of a speaker a certain constant coloring, and they play an important role in sociophonetics (Pratt, 2020). However, they also vary with the activity and potency features of emotions (Waaramaa et al., 2010) and are, additionally, very susceptible to effects of signal compression (Rozborski, 2007). For these reasons, we included mean F1-F3.

Given that we measured six tonal and seven non-tonal parameters, the acoustic analysis was based on 13 prosodic parameters in total. All parameters were measured per sentence. There were no pauses within the individual sentences. So, each sentence corresponds to a major prosodic phrase or an inter-pausal unit (IPU). Consequently, no pauses had to be taken into account or excluded, e.g., when measuring the speaking rate.

In our within-subjects study design, a normalization of measurements was basically not required. Individual speakers and speaker gender were perfectly balanced. That is, each condition included the same number of male and female tokens from the same speakers. Yet, while the measurements of parameters like speaking rate, intensity and F1, F2, and F3 differ between male and female speakers on average by about 10% or less (Pausewang Gelfer and Young, 1997; Iseli et al., 2007; Simpson, 2009), f0 measurements differ by almost factor two between male and female speakers, i.e., about 10 semitones on average (Traunmüller and Eriksson, 1995; Iseli et al., 2007; Simpson, 2009; Andreeva et al., 2014; Pépiot, 2014). Therefore, to reduce the statistical noise in our data, we gender-normalized the measurements of all f0 parameters by downscaling the women's values per speaker and IPU to the men's level based on the recent reference values in Pépiot (2014). That is, all female f0 values were recalculated by subtracting 10.5 semitones from each measurement, thereby bringing the female values in line with those of male speakers (the mean gender-specific f0 difference in Pépiot, 2014, which corresponds well to the difference determined by Andreeva et al., 2014).

This normalization represents a proportional rescaling of the measurements that leaves all other differences and aspects of the data intact. It was to be expected that at least part of the compression effects of the codecs on emotion acoustics would manifest itself in these other differences, such as distortions and blurring of measured parameter values (in a way that is perceptually relevant, see Siegert and Niebuhr, 2021b). For this reason we also refrained from, for example, a z-score transformation of measurements for all codec conditions of each sentence. Such a normalization could mask compression effects of the codecs on the emotion acoustics.

2.4. Statistics

The statistical analysis of the data was performed separately for the six tonal and seven non-tonal parameters and based on two types of tests. One of them was a multivariate analysis of covariance, MANCOVA. Two MANCOVAs tested separately for the sets of tonal and non-tonal parameters the effects of codec compression on the individual emotions, taking the individual four speakers into account. That is, Emotion (7 levels) and Codec (6 levels) were the two fixed factors (independent variables). The tonal and non-tonal measurement sets were the dependent variables. Sentence (in terms of its word count) was included as a covariate in the respective statistical model.

However, the key type of test in our analysis was a series of LDAs. We wanted to know whether the acoustic-prosodic parameters would lead to a significant discrimination of the seven emotion categories (including neutral) and whether and how much this discrimination performance deteriorates in combination with which compression codec. Therefore, the results of the LDAs are reported first in the Results section below. Note that, unlike analyzes of variance with equal sample sizes per condition, LDAs are not considered to be robust to data distortions, as we assume them to be more or less strongly created by the compression codecs. What is a problem in other studies—because the goal is usually to achieve the best possible discrimination, irrespective of data properties—is desirable in our study. We want a procedure, which provides us with sensitive insights into how the codecs worsen acoustic emotion discrimination. The LDA gives us these insights. Apart from that, the LDA is "currently the most used method for emotion classification in vocal stimuli" (Lausen and Hammerschmidt, 2020:3) and, thus, makes our results comparable to those of previous studies, not all of which used normalized or z-transformed data either (e.g., Hammerschmidt and Jürgens, 2007).

3. Results

Given our primary, emotion-oriented research questions, the key statistical analyzes were a series of LDAs. We wanted to know whether the acoustic-prosodic parameters would

lead to a significant discrimination of the seven emotion categories (including neutral) and whether and how much this discrimination performance deteriorates in combination with which compression codec.

Figure 3 shows the key result of this series of predictive models: Relative to the WAV reference files, there are clear codec-specific effects on the discrimination performances of the LDAs. However, these effects differ fundamentally for the tonal and non-tonal parameters. For the tonal parameters, the discrimination performance does not deteriorate by signal compression. Rather, the opposite is true: it gets better. The WAV files achieve an overall performance of 50.0% for the seven emotion categories (chance level being 1/7 or 14.3%, because we have included 7 emotional categories and all are represented by the same number of sentences. Thus, 50%, although it often refers to a chance level performance, actually means a correct performance rate three times better than chance level in our study). When using the MP3 and SPEEX codecs, this overall performance increases to 76.9% and 69.2% respectively. An exactly opposite pattern emerged for the non-tonal parameters. At 88.5%, the overall discrimination performance of the seven emotion categories was highest for the uncompressed WAV files. This reference performance then gradually drops across the codecs, is only 80.8% for MP3 and reaches the low point of 73.1% for SPEEX. Note, however, that all of these performances are generally higher than for the six tonal parameters.

In addition, Table 2 shows that the tonal and non-tonal LDAs also differ with respect to the emotions they could only discriminate poorly. For the tonal parameters, joy and sadness as well as the neutral rendering of the sentences were confused particularly often with other emotional categories. In the non-tonal LDAs, high confusion rates also occurred for joy and sadness, but additionally concerned anger, fear, and disgust. The neutral sentences were hardly confused with other, emotionally produced sentence. Furthermore, Table 2 shows that the top-three predictors the overall discrimination performance relied on were remarkably constant within each of the LDA series, the tonal and the non-tonal one. In the tonal LDA series, the f0-range parameters hardly played a role; what was primarily important were mean f0 level and the frequency scaling of the local minimum and maximum f0 values. In the non-tonal LDA series, the speaking rate did not play a major role. Instead, prediction or discrimination performances mainly

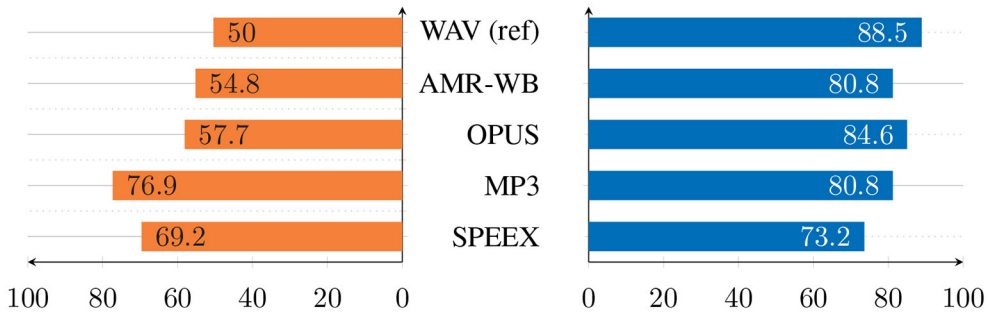


FIGURE 3
Overall emotion-discrimination performances (in %) of LDA series conducted per codec for the sets of tonal **(Left)** and non-tonal **(Right)** parameters. *N* = 26 per percentage.

TABLE 2 Summary of key results of the LDAs conducted per codec for the sets of tonal (top) and non-tonal (bottom) parameters; J, joy; B, boredom; S, sadness; N, neutral; F, fear; D, disgust; A, anger.

Condition	Wilks-λ	Test statistics	Top three predictors	Least predictable emo.
WAV	0.069	$\chi^2[36] = 51.223, p = 0.05$	Mean f0, Min f0, f0 Std	J, B, S
AMR-WB	0.062	$\chi^2[36] = 58.670, p = 0.03$	Mean f0, Min f0, f0 Max	N, J, S
OPUS	0.047	$\chi^2[36] = 50.100, p = 0.04$	Mean f0, Min f0, f0 Std	J, F, S
MP3	0.034	$\chi^2[36] = 62.703, p = 0.004$	f0 Min, f0 Max, Span	N, J, B
SPEEX	0.033	$\chi^2[36] = 63.235, p = 0.003$	f0 Min, p90 Range, f0 Std	N, D, S
WAV	0.017	$\chi^2[42] = 73.461, p = 0.002$	Mean F3, Mean int, H1-A3	J, A, S
AMR-WB	0.009	$\chi^2[42] = 83.871, p < 0.001$	Mean F3, Mean int, H1-H2	N, D, S
OPUS	0.011	$\chi^2[42] = 81.384, p < 0.001$	Mean F3, Mean int, Mean F1	F, J, A
MP3	0.013	$\chi^2[42] = 78.722, p < 0.001$	Mean F3, Mean int, H1-A3	N, J, S
SPEEX	0.024	$\chi^2[42] = 67.424, p = 0.008$	Mean F3, Mean F1, H1-A3	F, D, A

Blue highlighted table cells mark the tonal, red the non-tonal parameter results.

TABLE 3 Summary of the two MANCOVAs.

Main effect	non-tonal	df1|df2	F	p	η_p^2	Tonal	df1|df2	F	p	η_p^2
Emotion	H1-H2	6|112	15.113	<0.001	0.447	f0 min	6|112	24.897	<0.001	0.572
	H1-A3	6|112	33.105	<0.001	0.639	f0 max	6|112	21.630	<0.001	0.537
	Rate	6|112	20.251	<0.001	0.520	key	6|112	41.119	<0.001	0.688
	Intensity	6|112	23.067	<0.001	0.553	span	6|112	7.710	<0.001	0.292
	meanF1	6|112	4.295	0.001	0.187	std	6|112	20.721	<0.001	0.526
	meanF2	6|112	4.550	<0.001	0.196	Range10-90	6|112	7.422	<0.001	0.284
	meanF3	6|112	36.345	<0.001	0.661					
Codec	H1-H2	5|112	7.803	<0.001	0.258	f0 min	5|112	0.128	0.986	0.006
	H1-A3	5|112	5.960	<0.001	0.210	f0 max	5|112	0.137	0.983	0.006
	Rate	5|112	1.076	0.378	0.046	key	5|112	0.107	0.991	0.005
	Intensity	5|112	11.254	<0.001	0.334	span	5|112	0.544	0.743	0.024
	meanF1	5|112	1.231	0.299	0.052	f0 std	5|112	0.849	0.518	0.037
	meanF2	5|112	2.708	0.024	0.108	Range10-90	5|112	3.363	0.007	0.131
	meanF3	5|112	12.993	<0.001	0.367					

Effect sizes are provided in terms of partial eta squared. Interaction effects were all not significant and are thus omitted in the summary.
Blue highlighted table cells mark the tonal, red the non-tonal parameter results.

relied on the voice-quality parameter H1-A3, the third formant F3, and the mean intensity.

Besides the two LDA series, we also calculated multivariate ANCOVAs, again separately for the two sets of tonal and non-tonal parameters. Sentence was integrated into the analyzes as a covariate. Emotion (7 levels) and Codec (6 levels) represented the two between-subjects fixed factors. Table 3 summarizes the relevant results of the two MANCOVAs. First, note that emotion resulted in main effects for almost all acoustic-prosodic parameters. That is, the seven emotional categories all differed clearly from each other in terms of the measured parameter set. It is beyond the scope of this study to present acoustic-prosodic profiles of all emotions involved, i.e., to detail how each emotion was characterized prosodically and to compare these characteristics to those carved out in previous studies. It is also well known from these previous studies how basic emotions differ acoustically (Mozziconacci, 2001; Gobl and Chasaide, 2003; Bänziger and Scherer, 2005); in fact, some comparative analyzes were also carried out for the EMODB corpus used here (Raju et al., 2018). However, note that two consistent patterns emerged across all 6 + 7 parameters. Firstly, the most extreme values along each parameter were always associated with the two emotions joy and anger on the one hand and boredom and sadness on the other, see, for example, Figure 4A for the emotion-specific f0 mimima. In contrast, there was, secondly, one single emotion whose measured values consistently fell in the middle of the acoustic range of all emotions. That is, one emotion did not stand out in any acoustic direction: disgust.

Unlike for the fixed factor Emotion, Table 3 shows hardly any significant main effects of Codec on the acoustic parameters, especially not on the tonal ones. Only for the p90 range, there was a significant main effect of Codec. For f0 std, the main effect approached significance. A closer inspection of the data shows that both effects are due to a fanning-out of the measured values. That is, different emotions are influenced by compression in different directions. For example, the f0 values of joy tend to increase, whereas those of anger tend to decrease. Indications in these directions can also be seen for the f0 minima in Figure 4A. The fanning-out increases the acoustic distinctiveness of the emotions under compression and, thus, also f0 range or variability measures, especially in connection with MP3 and SPEEX. For the set of non-tonal parameters, there are a few significant main effects of Codec. This applies in particular to those measures that also had high prediction performances in the LDAs: mean F3, mean intensity, H1-A3, and H1-H2. The main effects are due to the fact that the values drop and merge to various degrees under speech compression. This applies even to something as popular as MP3 compression, but is especially pronounced under SPEEX compression. The latter is shown in Figure 4B, using the mean F3 as an example.

Note that there are no significant interactions of Emotion and Codec. The covariate Sentence yielded no significant effect either.

4. Discussion and conclusion

4.1. Interpretation of the findings

Siegert and Niebuhr (2021a) found that strong speech signal compression has significant effects on the perception of speaker charisma and related speaker traits. The perceived speaker charisma is significantly reduced, especially in connection with MP3 and SPEEX compression. Siegert and Niebuhr (2021b) then searched for the acoustic foundations of these perceptual changes and were able to determine that (a) codec compression systematically distorts the automatic measurements of charisma-relevant acoustic-prosodic parameters, that (b) these distortions

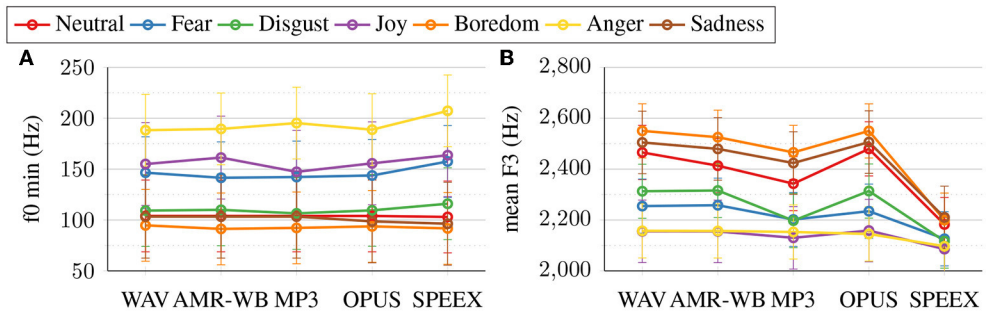


FIGURE 4
Estimated marginal means regarding Emotion*Codec interactions for the tonal parameter f0 min **(A)** and the non-tonal parameter mean F3 **(B)**. Each datapoint *N* = 26.

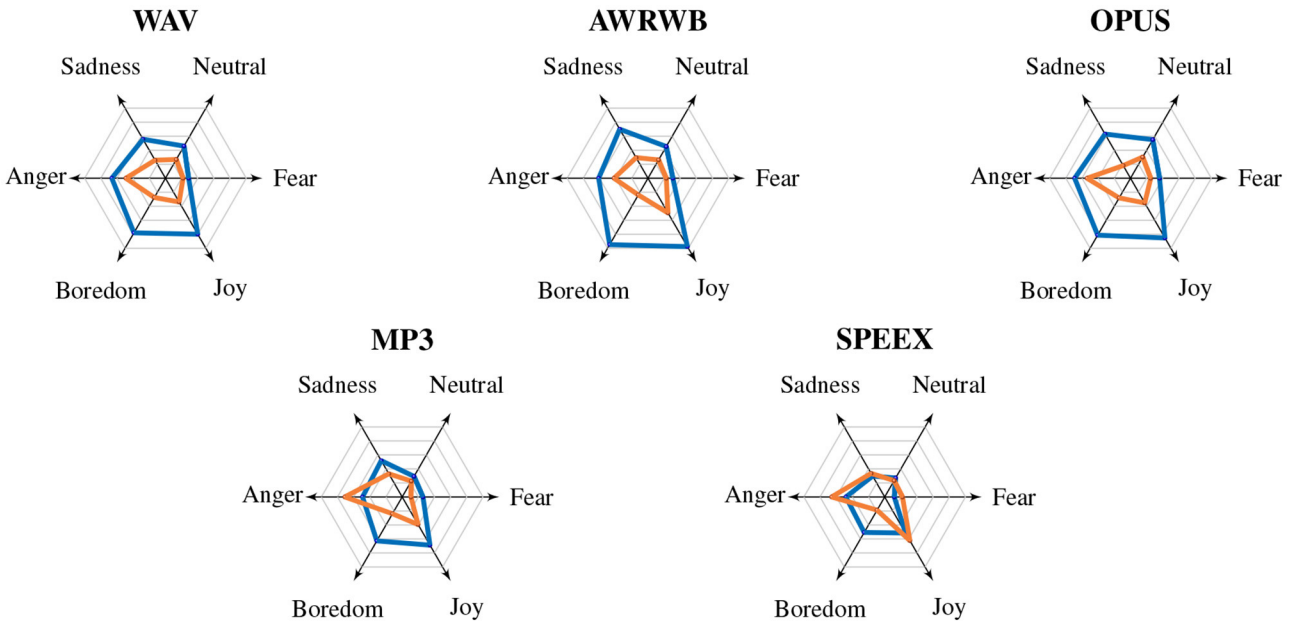


FIGURE 5
Spiderweb diagram for non-tonal (–) and tonal (–) moduli of the group centroids for the significant discriminant function(s) of the LDAs.

are less strong for f0-based parameters than for spectral ones and less strong for holistic parameters than for ones that refer to local measurements, and that (c) the magnitude of codec-induced parameter changes correlates well with magnitude the lowering of perceived speaker charisma caused by these changes.

Niebuhr and Siegert (2022) extended this line of research to emotion perception using the same speech corpus (EMO-DB) and the same rating scales; firstly, because emotions are a key component of charisma (especially with regard to the degree of arousal) and, secondly, because the question to what extent codecs distort emotion recognition is of particular relevance to the everyday life of language users. Niebuhr and Siegert found that the seven emotion categories tested (incl. neutral) formed individual perception profiles along charisma-related rating scales—that compression blurred these emotion-specific profiles to codec-specific degrees, up to the point that they completely merged. That is, emotional speaker states were no longer perceptually distinct. Analogous to Siegert and Niebuhr (2021b), the task of the present acoustic analysis was to look for the foundations of these perceptual findings. If, to what degree, and along which parameters does an acoustic analysis reflect these perceptual findings? The conducted acoustic analysis relied on the same prosodic parameters as in Siegert and Niebuhr (2021b), i.e., temporal parameters were not included as they have already proved irrelevant in Siegert and Niebuhr (2021b).

Three questions were addressed. The first question concerned the qualitative nature of the connection between the perceptual distortions and their prosodic correlates: How does speech signal compression affect the acoustics of different emotions? The second question was conceptual in nature and asked which further insights can be gained from the above connection about the acoustic correlates of charisma and the relation of charisma and (different) emotions: How is emotion acoustics related to listener ratings of perceived speaker charisma and related traits? The third question

was of a more specific, practical nature and asked which concrete acoustic patterns could be associated to the distortions in emotion perception: Does the merging of emotions along the charisma scales have an acoustic correspondence, and if so, along which prosodic parameters? The results of the present acoustic analysis provide the following answers to these three questions.

How does speech signal compression affect the acoustics of different emotions? The most important answer to this question is that the emotional acoustics are significantly and in some conditions to a considerable degree influenced by signal compression. The main effects of Codec in the MANCOVAs show this very clearly. The nature of this influence is complex. For f0 parameters, we find a fanning out of the measured values in emotion-specific directions. The (automatically measured) values can deviate by up to 20% from the values of the WAV reference condition. In particular, measures such as the (p90) f0 range or the f0 variability (std. dev.) are affected, which relate multiple values to one another, including local landmark values such as the f0 minimum. In the WAV condition, for example, the f0 variability values of joy and anger are only about 4 Hz apart (48.5 Hz vs. 52.7 Hz). Under MP3 compression, this difference increases to 22 Hz (40.3 Hz vs. 62.9 Hz)—and under SPEEX compression even to 29 Hz (38.2 Hz vs. 67.6 Hz). Figure 4A illustrates that this effect is also there for the f0 minimum, albeit less pronounced. For spectral measures, the effects of signal compression are qualitatively different. First, they are more numerous and stronger than for f0 (i.e. there are more significant main effects of Codec). Thus, this study replicates the findings of Siegert and Niebuhr (2021b), who also found that f0 measures are less susceptible to signal-compression effects than non-tonal measures. That f0 parameters are more robust against codec compression than spectral parameters is moreover in accord with van Son (2005) and Bollepalli (2013), whose studies were concerned with MP3 and Ogg Vorbis compressions. Our results suggest that this difference between the two different types of measures generally applies to codec compression. Second, the effects of compression on non-tonal parameters are more homogeneous across emotions. That is, the codecs affect the levels of measured values, but they do this similarly for all emotions, so that the acoustic distances between the emotions are more strongly preserved. A notable exception is SPEEX. This codec leads to an extensive merging of the non-tonal measurements across all emotions, as is exemplified in Figure 4B for the mean F3. The described differences are responsible for the fact that the LDA-based emotion discrimination gets almost 20% *better* under compression for tonal but more than 15% *worse* for non-tonal parameters.

The second question was: *How is emotion acoustics related to listener ratings of perceived speaker charisma and related traits?* Our findings provide a number of insights related to this question, based on clear parallels between the present results and the listener ratings in Niebuhr and Siegert (2022). In the latter study, joy and anger emerged as similarly charismatic emotions, and in fact these two emotions were also the most frequently confused ones in our acoustics-based LDAs, particularly in the WAV condition. Joy and anger stand out from the other emotion categories (incl. neutral) by the following characteristics: low values of H1-H2 and H1-A3, i.e. a shallow spectral slope, a high intensity level, high F1 values in combination with low F2 and F3 values, as well as the highest/largest f0 values of all emotions, including those of variability and range. These acoustic profiles are broadly consistent with those described as charisma-inducing in Niebuhr and Silber-Varod (2021). By far the most uncharismatic emotion in the perception study by Niebuhr and Siegert (2022) was disgust, ahead of boredom and sadness. In the present acoustic analysis, disgust emerged as the emotion whose acoustics was in the middle of all emotions. Expressing disgust meant to be neither particularly fast nor slow, neither high-pitched nor low-pitched, neither loud nor soft. Disgust meant relative mediocrity. This seems to imply that, not to stand out means not to be charisma-inducing. Furthermore, Niebuhr & Siegert reported that no significant emotion differentiation based on the charisma-related scales was possible under SPEEX. The present data show that the emotion discrimination of the LDAs improved for the f0 parameters, but worsened for the temporal and spectral parameters. All in all, these parallels between the present findings and those of Niebuhr & Siegert allow the following conclusions: In the tonal domain, charisma perception is essentially based on raised, more extensive, and more frequent (i.e., more variable) f0 movements. In the non-tonal domain, an essential ingredient of charismatic speech is an increased "vocal effort", which (in addition to a raised f0) manifests itself in higher F1 and H1-A3 values (Liénard, 1999). The speaker's voice is loud, sounds resonant and is tense (negative H1-H2 values) than rather than breathy and thin. The speaker sounds big, i.e., the acoustic body size is high (lower F2 and F3 values; González, 2004). Moreover, these spectral aspects of prosody seem to carry more weight for perceived speaker charisma than the tonal aspects (in view of the comparisons between LDAs and the listener ratings in Niebuhr and Siegert 2022). This conclusion also fits with the latest data from Niehof and Niebuhr (2022), who concluded on the basis of their analyzes that a good tonal performance is required for investor pitches to win competitions—but it is a good spectral (and temporal) prosodic performance that decides about whether investor pitches loose competitions.

The third question was: *Does the merging of emotions along the charisma scales have an acoustic correspondence, and if so, along which prosodic parameters?* The answer to the second question already indicated that the answer to the third question is clearly positive. There are remarkable parallels between the codec-related merging of the emotion categories along the listener ratings in Niebuhr and Siegert (2022) on the one hand and the confusion and merging of emotion categories in the present LDAs on the other hand. Joy and anger were rated similarly charismatic in Niebuhr and Siegert (2022) and were particularly often confused in the present LDAs. The opposite is true for boredom and anger. These two emotions received very different charisma ratings in Niebuhr and Siegert (2022) and were not confused at all in the present LDAs. The charisma-related emotion ratings in Niebuhr and Siegert (2022) merged under SPEEX compression, and the same was true here to a large extent for the spectral and temporal prosodic parameters. Accordingly, the merging of the listener ratings in Niebuhr and Siegert (2022) is reflected here in the spectral dimensions—most

strongly in the following three: intensity (RMS), mean F2 and mean F3.

To conclude the discussion of the results, the displays in Figure 5 provide a graphical summary of the findings. Shown are the largest moduli of the group centroids for the significant discriminant function(s) of the LDA—separately for each individual emotion. The centroids are mulivariate averages. They quantify how much a group—in this case an emotion—differs in standard deviation units from the zero mean of the discriminant function. They are thus a measure of the separability of the individual emotions from one another in the multivariate acoustic parameter space. As in Figure 3, the tonal and non-tonal values are differently color coded. At the level of the individual codecs it can be seen in Figure 5 that distinct shapes and area sizes emerge from the moduli of all emotions. The area size directly reflects the overall degree to which a codec can separate the individual emotions. In addition, how complementary (or non-complementary) the areas of the non-tonal and tonal moduli are in the value space contributes to the overall separability, too—as does a more circular shape as a sign of a more homogeneous discrimination performance across emotions. Consistent with the percentages in Figure 3, we see in Figure 5 that for WAV, as well as for OPUS and AMR-WB, the non-tonal areas are significantly larger than the tonal area. For WAV and OPUS, the areas are more circular, though, than the rather angular area of AMR-WB. We can also see that the non-tonal area for MP3 is already smaller than for the other three codecs, and in the case of SPEEX compression it shrinks considerably again. Simultaneously, SPEEX yielded the largest tonal area of all codecs. It is even larger than SPEEX's non-tonal area, albeit the least circular of all codecs. MP3 compression created a similarly large tonal area as SPEEX, but more circularly, which led to the overall highest tonal separation performance of all codecs (cf. Figure 3). For the WAV reference condition and the other two codecs, OPUS and AMR-WB, the tonal area is so much smaller than for SPEEX and MP3 that it lies entirely within the non-tonal area. Furthermore, the displays in Figure 5 also reflect the intermediate acoustic expression of disgust (in the form of consistently low group centroid moduli) and the generally more pronounced prosodies of anger and joy, with group centroid moduli for the latter emotion being higher than for the former, which reflects that anger was more often confused with joy than joy with anger.

At the beginning of the paper and in its title, we asked whether there is a digital "flat affect" for emotional prosody. At the end of the results' discussion, the answer to this question is a conditional "yes". Overall, this digital "flat affect" seems to apply more strongly to perception, for which emotional mergers were even more pronounced than in acoustics, and in the latter domain the "flat affect" did not apply to all codecs and all prosodic parameters. The non-tonal parameters seemed to be affected more than the tonal ones. In any case, it is particularly relevant for the nature of the "flat affect" that high-arousal emotions are more likely to be affected by acoustical codec distortions. For this reason, we can conclude even in view of our study's limitations (see Section 4.3) that a digital "flat affect" cannot be categorically ruled out in the everyday use of digital communication devices.

4.2. Toward explaining the compression effects

Finding technical explanations for the codec-related distortions of the acoustic-prosodic measurements goes beyond the aims of this study. Our goal was to test the audio files for the existence of such distortions and to examine the extent to which they match the results pattern of emotion perception from Niebuhr and Siegert (2022). Explaining the distortions technically would require correcting or manipulating the assumed origins of the distortions in the codecs so as to compare their new acoustic outputs and measurements with those of the original codec versions. This is an engineering question and no longer a socio- or psycho-phonetic question. However, paving the way to this engineering question, we offer below some assumptions about which basic mechanisms in the codecs could be responsible for the distortions.

First, we suspect the frame size of the codecs to be one of the three main causes, especially if the frame size is large and/or fixed. One of the defining characteristics of emotional (and also charismatic) speech vs. matter-of-fact speech is its altered signal dynamics (Banse and Scherer, 1996; Mozziconacci, 2001; Gobl and Chasaide, 2003; Niebuhr and Silber-Varod, 2021 for charisma). For higher arousal emotions like anger and joy, prosodic changes take place faster, and they involve larger contrasts, i.e., the used value range increases and syntagmatic differences are realized more pronounced, such as that between stressed and unstressed syllables. The opposite is true for low-arousal, low-dynamics emotions like sadness. It is plausible to assume that codecs can only inadequately represent altered dynamics, which deviate greatly from the emotionally neutral standard, if the frame size is not appropriate. A too large frame size means, for high arousal emotions, more changes than the codec can handle. Conversely, if the frame size is too small and the signal dynamics is too low, relevant prosodic changes could be overlooked by the codec. This would explain why emotions with particularly high/low arousal levels were (in terms of measurement distortions) acoustically changed the most by the codecs, either in the direction of better or worse discrimination to neighboring emotions in the LDA series. It would additionally explain why the MP3 and SPEEX compressions with their relatively large, fixed frame sizes of 24 ms or 30 ms (cf. Table 1) caused the greatest acoustic distortions. Furthermore, it would also explain why speaking rate, i.e., the measurement of defined intensity peaks and their frequency of occurrence per second, was not a major emotion discrimination factor for the compressed audios, although the general relevance of this factor is well documented (Banse and Scherer, 1996; Mozziconacci, 2001). Intensity peaks have a significantly higher dynamics than, for example, f0 or formant transitions and are therefore more susceptible to measurement errors and more variability within and between the emotion categories. In fact, for instance for SPEEX, the average speaking rate across all emotions is about 10% lower, i.e., intensity peaks got lost in the compression, while the variability of the speaking rate across all emotions increased.

A second reason, directly related to the first, we suspect is the size and adaptivity of a codec's codebooks. A codebook contains the range of quantization vectors that a codec is able

to use. This applies in particular to the source signal. In the case of emotional speech, this source signal deviates partially or continuously from the modal voice signal of neutral utterances. This, in connection with the different f0 dynamics, could exceed the possibilities of some codebooks, such as SPEEX, whose codebook is not adaptive but hard-coded. The two codecs AMRWB and OPUS have larger and more adaptive codebooks. Only MP3 is based on a completely different compression method that does not require codebooks. It is striking in this context that MP3 was the only compression method that had the f0 range among the top-three tonal predictors and the only one whose non-tonal top-three predictors were identical to those of the WAV reference condition.

The third and perhaps most obvious reason is the range of spectral frequencies that the codecs took into account. We concluded here in line with, for example, Gobl and Chasaide (2003) that emotions are, in both dimensions arousal and valence, significantly defined by (differences in) voice quality, i.e., by (differences in) spectral energy distributions. Since the codecs took different spectral frequency ranges into account, it was therefore to be expected that this would cause perceptual distortions that match acoustically measured parameter distortions. SPEEX only covered the lower 4 kHz of the speech signal, AMR-WB reached up to 6.4 kHz, and OPUS included spectral frequencies until about 8 kHz. This order corresponds to the performance that the codecs achieved in terms of perceptual and acoustic discrimination relative to the WAV reference condition. Furthermore, it is consistent with this frequency-range explanation that emotions with a tenser voice and hence with a shallower spectral slope compared to neutral utterances—i.e., fear, anger, disgust—lost more of their acoustic discriminability through codec compression than other emotions.

4.3. Limitations and outlook

The present study had some limitations. This includes, among other things, that (1) the acoustic analysis was based on comparatively few speakers, that (2) the speech material was based on enacted emotions, and that (3) only very high compression rates of selected codecs were examined. For the present question, these limitations were acceptable. As we argued in 2.1, the aim here was to investigate the general nature of acoustic distortion effects caused by some technical signal-processing algorithms and not, for example, to separate meaningful prosodic patterns from variable human speech behavior. For the success of this investigation, it was the category diversity of the speech material that mattered and not its ecological validity—also because Scherer (2013) showed that the prosody of enacted emotions can basically be comparable with that of authentic emotional field recordings. It was also useful and reasonable here to begin this investigation with looking at a very high compression levels, for which the strongest effects could be expected. In addition, note that all limitations were not only acceptable—they were mandatory, because as we wanted to project the measured acoustic distortion effects onto the perception results of Niebuhr and Siegert (2022), which required using the same speech data, compression methods, and compression levels as in the latter study.

Although this projection was successful in that it revealed parallels between acoustic distortions and the listener ratings in Niebuhr and Siegert (2022), the above points (1)-(3) remain relevant limitations of this study. Regardless of these limitations, there can be no doubt in view of Section 4.1 that we were able to show here that codec compression of speech signals not only influences the naturalness and intelligibility of matter-of-fact utterances. It also influences the acoustics of emotions and—as shown by Niebuhr and Siegert (2022)—the identification of emotions in human speech. In this respect, the present data represent a call-to-action to go beyond phoneme and word recognition as performance criteria in researching and developing speech-signal compression procedures. Everyday conversations include a lot more than conveying and requesting propositional content, but this fact is currently not sufficiently reflected in speech compression/transmission technology.

When it comes to the details and diversification of this general call-to-action, the above limitations become relevant. Future research must therefore focus on comparatively examining the acoustics and perception of emotions—and higher-level concepts such as speaker charisma—taking into account the abundance of emotional variation and the availability of other codec variants and compression levels.

The first task of such follow-up studies should be to apply the codec variants and compression levels tested here to a wider range of speech material, i.e., to more speakers and more utterances. An extended speaker sample means additional voice-quality and f0-level conditions, and it also opens up the possibility of a separate investigation of speaker-sex effects (which are already well documented for charisma and uncompressed emotions, cf. Parkins, 2012; Niehof and Niebuhr, 2022). Testing more utterances means looking at a more diverse interplay of segmental and prosodic phenomena and, thus, at a richer and more variable prosody. In addition, follow-up studies could also test other languages and/or in-situ recorded emotions. For example, we assumed here that emotions which are less clearly pronounced (compared to enacted ones) would exacerbate compression-induced the acoustic overlap and, thus, the perceptual confusion of emotional categories. In principle, however, it is also possible that less strong emotions reduce the signal distortion caused by codecs as less strong typically means less extreme prosodic dynamics, in this way preserving the identification of emotional categories. Overall, our speaker and utterance samples are clearly too small to draw differentiated conclusions about exactly how, where, and under what conditions codecs weaken emotional signals.

A second task of follow-up studies, which is orthogonal to the first one, is to vary the codecs themselves, either with respect to the levels, versions, and types of codec compression or by iteratively modifying and testing the compression algorithms or procedures themselves. The first type of variation is relevant for everyday life, for example, in that it will allow determining compression thresholds up to which the distortion of the emotion

acoustics and the identification of emotional categories are still acceptable. The targeted modification of the compression algorithms or procedures, on the other hand, is of conceptual relevance in that will allow revealing sources of distortions, to understand them and, if possible, to minimize them.

A third task of follow-up studies could ultimately be to use speech compression not as a research topic but as a research tool, for example, in order to gain insights into the importance of tonal and non-tonal prosodic cues to emotions, charisma, and related phenomena. The opposite LDA performances found here for tonal and non-tonal prosodic parameters, in combination with previous perception results that more closely resemble the non-tonal LDA performances (Niebuhr and Siegert, 2022) show the potential of this approach. This is all the more relevant since previous research into phenomena such as emotions and charisma has often focused on tonal parameters, while the factor voice-quality is currently gaining attention.

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found at: http://emodb.bilderbar.info/index-1024.html.

Author contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

Conflict of interest

ON is the CEO and founder of the speech technology company AllGoodSpeakers ApS. Please visit https://oliverniebuhr.com for a conflict-of-interest statement.

The remaining author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Ahern, S. (2020). *Acoustical Design of Concert Halls and Theatres: A Personal Account, 3rd Edn*. London: Routledge.

Albahri, A., Lech, M., and Cheng, E. (2016). Effect of speech compression on the automatic recognition of emotions. *Int. J. Signal Process. Syst.* 4, 55–61. doi: 10.12720/ijsps.4.1.55-61

Allison, T. H., Warnick, B. J., Davis, B. C., and Cardon, M. S. (2022). Can you hear me now? engendering passion and preparedness perceptions with vocal expressions in crowdfunding pitches. *J. Bus. Venturing* 37, 106193. doi: 10.1016/j.jbusvent.2022.106193

Alpert, M., and Anderson, L. T. (1977). Imagery mediation of vocal emphasis in flat affect. *Arch. Gen. Psychiatry* 34, 208–212. doi: 10.1001/archpsyc.1977.01770140098011

Alpert, M., Rosenberg, S. D., Pouget, E. R., and Shaw, R. J. (2000). Prosody and lexical accuracy in flat affect schizophrenia. *Psychiatry Res.* 97, 107–118. doi: 10.1016/S0165-1781(00)00231-6

Amari, Y., Okada, S., Matsumoto, M., Sadamitsu, K., and Nakamoto, A. (2021). "Multimodal analysis of client persuasion in consulting interactions: toward understanding successful consulting," in *Social Computing and Social Media: Applications in Marketing, Learning, and Health*, ed G. Meiselwitz (Cham: Springer International Publishing), 29–40.

Andreeva, B., Demenko, G., Wolska, M., Möbius, B., Zimmerer, F., Jügler, J., et al. (2014). Comparison of pitch range and pitch variation in slavic and germanic languages. *Proc. Speech Prosody* 2014, 776–780. doi: 10.21437/SpeechProsody.2014-144

Antonakis, J., Bastardoz, N., Jacquart, P., and Shamir, B. (2016). Charisma: an ill-defined and ill-measured gift. *Annu. Rev. Organ. Psychol. Organ. Behav.* 3, 293–319. doi: 10.1146/annurev-orgpsych-041015-062305

Antonakis, J., Fenley, M., and Liechti, S. (2011). Can charisma be taught? tests of two interventions. *Acad. Manag. Learn. Educ.* 10, 374–396. doi: 10.5465/amle.2010.0012

Arvaniti, A. (2020). *The Phonetics of Prosody*. Oxford: Oxford University Press.

Audibert, N., Aubergé, V., and Rilliard, A. (2010). "Prosodic correlates of acted vs. spontaneous discrimination of expressive speech: a pilot study," in *Proceedings of the 5th International Conference of Speech Prosody* (Chicago, IL), 1–4.

Banse, R., and Scherer, K. R. (1996). Acoustic profiles in vocal emotion expression. *J. Pers. Soc. Psychol.* 70, 614–636. doi: 10.1037/0022-3514.70.3.614

Bänziger, T., and Scherer, K. R. (2005). The role of intonation in emotional expressions. *Speech Commun.* 46, 252–267. doi: 10.1016/j.specom.2005.02.016

Barnes, J., Veilleux, N., Brugos, A., and Shattuck-Hufnagel, S. (2012). Tonal center of gravity: a global approach to tonal implementation in a level-based intonational phonology. *Lab. Phonol.* 3, 337–383. doi: 10.1515/lp-2012-0017

Bass, B. (1990). *Bass and Stogdill's Handbook of leadership: Theory, Research, and Managerial Applications, 3rd Edn*. New York, NY: Free Press.

Bhangale, K., and Mohanaprasad, K. (2021). A review on speech processing using machine learning paradigm. *Int. J. Speech Technol.* 24, 367–388. doi: 10.1007/s10772-021-09808-0

Bollepalli, B. (2013). *Effect of MPEG audio compression on HMM-based speech synthesis*. Internship report, Aalto University, Finlad.

Bono, J. E., and Ilies, R. (2006). Charisma, positive emotions and mood contagion. *Leadership Q.* 17, 317–334. doi: 10.1016/j.leaqua.2006.04.008

Brandenburg, K. (1999). "MP3 and AAC explained," in *17th AES International Conference: High-Quality Audio Coding* (Florence).

Bundesnetzagentur (2022). "Karte," in *Breitband-Monitor*. Available online at: https://gigabitgrundbuch.bund.de/GIGA/DE/Funklochkarte/start.html (accessed March 9, 2023).

Burkhardt, F., Paeschke, A., Rolfes, M., Sendlmeier, W., and Weiss, B. (2005). "A database of german emotional speech," in *Proceedings of the INTERSPEECH-2005* (Lissabon), 1517–1520.

Byrne, C., and Foulkes, P. (2004). The 'mobile phone effect' on vowel formants. *Int. J. Speech Lang. Law* 11, 83–102. doi: 10.1558/sll.2004.11.1.83

Campbell, N., and Mokhtari, P. (2003). "Voice quality: the 4th prosodic dimension," in *Proceedings of 15th International Congress of Phonetic Sciences* (Barcelona), 2417–2420.

Caviglione, L. (2015). A first look at traffic patterns of siri. *Trans. Emerg. Telecommun. Technol.* 26, 664–669. doi: 10.1002/ett.2697

Chavarría, S., Yoon, T.-J., Cole, J., and Hasegawa-Johnson, M. (2004). "Acoustic differentiation of ip and IP boundary levels: comparison of l- and l-l% in the switchboard corpus," in *Proceedings of 2nd International Conference of Speech Prosody* (Nara), 333–336.

Chebat, J.-C., Hedhli, K. E., Gélinas-Chebat, C., and Boivin, R. (2007). Voice and persuasion in a banking telemarketing context. *Percept. Mot. Skills* 104, 419–437. doi: 10.2466/pms.104.2.419-437

Chen, L., Feng, G., Joe, J., Leong, C. W., Kitchen, C., and Lee, C. M. (2014). "Towards automated assessment of public speaking skills using multimodal cues," in *Proceedings of the 16th International Conference on Multimodal Interaction, ICMI '14* (New York, NY: ACM), 200–203. doi: 10.1145/2663204.2663265

Damen, F., Van Knippenberg, D., and Van Knippenberg, B. (2008). Leader affective displays and attributions of charisma: the role of arousal. *J. Appl. Soc. Psychol.* 38, 2594–2614. doi: 10.1111/j.1559-1816.2008.00405.x

Davis, B. C., Hmieleski, K. M., Webb, J. W., and Coombs, J. E. (2017). Funders' positive affective reactions to entrepreneurs' crowdfunding pitches: the influence of perceived product creativity and entrepreneurial passion. *J. Bus. Venturing* 32, 90–106. doi: 10.1016/j.jbusvent.2016.10.006

D'Errico, F., Niebuhr, O., and Poggi, I. (2019). "Humble voices in political communication: a speech analysis across two cultures," in *Computational Science and Its Applications-ICCSA 2019* (Cham: Springer), 361–374.

D'Errico, F., and Poggi, I. (2022). 'emotion that goes around comes around'. Political charisma and the emotions conveyed through social media. *Acta Polytechn. Hungarica* 19, 173–194. doi: 10.12700/APH.19.5.2022.5.9

Evensen, J., Røssberg, J. I., Barder, H., Haahr, U., Hegelstad, W. T. V., Joa, I., et al. (2012). Flat affect and social functioning: a 10 year follow-up study of first episode psychosis patients. *Schizophr. Res.* 139, 99–104. doi: 10.1016/j.schres.2012.04.019

Feraru, S. M., Schuller, D., and Schuller, B. (2015). "Cross-language acoustic emotion recognition: an overview and some tendencies," in *2015 International Conference on Affective Computing and Intelligent Interaction (ACII)* (Xi'an), 125–131. doi: 10.1109/ACII.2015.7344561

Fischer, K., Niebuhr, O., and Asadi, A. (2022). "The voice of creativity: effects of pitch range in the voice of a robot facilitator," in *Elektronische Sprachsignalverarbeitung 2022, volume 81 of Studientexte zur Sprachkommunikation* (Sonderborg), 121–130.

Fischer, K., Niebuhr, O., Jensen, L. C., and Bodenhagen, L. (2019). Speech melody matters–how robots profit from using charismatic speech. *J. Hum. Robot Interact.* 9, 274. doi: 10.1145/3344274

Furnham, A., Chan, P. S., and Wilson, E. (2013). What to wear? the influence of attire on the perceived professionalism of dentists and lawyers. *J. Appl. Soc. Psychol.* 43, 1838–1850. doi: 10.1111/jasp.12136

Gobl, C., and Chasaide, A. N. (2003). The role of voice quality in communicating emotion, mood and attitude. *Speech Commun.* 40, 189–212. doi: 10.1016/S0167-6393(02)00082-1

González, J. (2004). Formant frequencies and body size of speaker: a weak relationship in adult humans. *J. Phon.* 32, 277–287. doi: 10.1016/S0095-4470(03)00049-4

Gooty, J., Connelly, S., Griffith, J., and Gupta, A. (2010). Leadership, affect and emotions: a state of the science review. *Leadership Q.* 21, 979–1004. doi: 10.1016/j.leaqua.2010.10.005

Gottheil, E., Paredes, A., Exline, R. V., and Winkelmayer, R. (1970). Communication of affect in schizophrenia. *Arch. Gen. Psychiatry* 22, 439–444. doi: 10.1001/archpsyc.1970.01740290055007

Goy, H., Pichora-Fuller, M. K., Singh, G., and Russo, F. A. (2016). Perception of emotional speech by listeners with hearing aids. *Can. Acoustics* 44, 182–183. Available online at: https://jcaa.caa-aca.ca/index.php/jcaa/article/view/2962

Grabo, A., Spisak, B. R., and van Vugt, M. (2017). Charisma as signal: an evolutionary perspective on charismatic leadership. *Leadership Q.* 28, 473–485. doi: 10.1016/j.leaqua.2017.05.001

Gregory, S. W., and Gallagher, T. J. (2002). Spectral analysis of candidates' nonverbal vocal communication: predicting u.s. presidential election outcomes. *Soc. Psychol. Q.* 85, 298–308. doi: 10.2307/3090125

Guillemin, B. J., and Watson, C. I. (2009). "Impact of the GSM AMR speech codec on formant information important to forensic speaker identification," in *Proceedings of IEEE ICASSP* (Instanbul: IEEE).

Hammerschmidt, K., and Jürgens, U. (2007). Acoustical correlates of affective prosody. *J. Voice* 21, 531–540. doi: 10.1016/j.jvoice.2006.03.002

Hazra, S., and Mateti, P. (2017). "Challenges in android forensics," in *Security in Computing and Communications*, eds S. M. Thampi, G. Martínez Pérez, C. B. Westphall, J. Hu, C. I. Fan, and F. Gómez Mármol (Singapore: Springer Singapore), 286–299.

Höbel-Müller, J., Siegert, I., Heinemann, R., Requardt, A. F., Tornow, M., and Wendemuth, A. (2019). "Analysis of the influence of different room acoustics on acoustic emotion features and emotion recognition performance," in *Tagungsband-DAGA 2019* (Rostock), 886–889.

House, R. (1977). *A 1976 Theory of Charismatic Leadership*, Carbondale: Southern Illinois; University Press.

Humphrey, R., Pollack, J., and Hawver, T. (2008). Leading with emotional labor. *J. Manag. Psychol.* 23, 151–168. doi: 10.1108/02683940810850790

Iseli, M., Shue, Y.-L., and Alwan, A. (2007). Age, sex, and vowel dependencies of acoustic measures related to the voice source. *J. Acoust. Soc. Am.* 121, 2283–2295. doi: 10.1121/1.2697522

ITU-T (1996). *Methods for subjective determination of transmission quality*. REC. P.800.

ITU-T (2003). *Wideband Coding of Speech at around 16 kbit/s using Adaptive Multi-Rate Wideband (AMR-WB)*. REC G.722.2.

ITU-T (2014). *Methods for objective and subjective assessment of speech quality (POLQA): Perceptual Objective Listening Quality Assessment*. REC P.863.

Jokisch, O., Iaroshenko, V., Maruschke, M., and Ding, H. (2018). "Influence of age, gender and sample duration on the charisma assessment of german speakers," in *Studientexte zur Sprachkommunikation: Elektronische Sprachsignalverarbeitung*, eds A. Berton, U. Haiber, and W. Minker (Dresden: TUD Press), 224–231.

Jokisch, O., Maruschke, M., Meszaros, M., and Iaroshenko, V. (2016). "Audio and speech quality survey of the opus codec in web real-time communication," in *Elektronische Sprachsignalverarbeitung 2016. Tagungsband der 27. Konferenz, volume 81 of Studientexte zur Sprachkommunikation*, ed O. Jokisch (Leipzig: TUD Press), 254–262.

Kao, Y.-H., and Lee, L.-S. (2006). "Feature analysis for emotion recognition from Mandarin speech considering the special characteristics of Chinese language," in *Proceedings of 17th International Interspeech conference* (Pittsburgh, PA), 1814–1817. doi: 10.21437/Interspeech.2006-501

Karabell, S. (2016). "Dressing like a leader: style tips for women in the spotlight," in *Forbes Magazin*. Available online at: https://www.forbes.com/sites/shelliekarabell/2016/01/16/dressing-like-a-leader-style-tips-for-women-in-the-spotlight/?sh=23e692412466 (accessed March 9, 2023).

Keating, C. F. (2011). "Channelling charisma through face and body status cues," in *Social Psychological Dynamics*, eds D. Chadee and A. Kostic (Jamaica: University of the West Indies Press), 93–111.

Kelman, H. C. (1961). Processes of opinion change. *Public Opin. Q.* 25, 57–78. doi: 10.1086/266996

Kienast, M., and Sendlmeier, W. F. (2000). "Acoustical analysis of spectral and temporal changes in emotional speech," in *Proceedings of ITRW on Speech and Emotion* (Newcastle), 92–97.

Kim, J., Cho, D., and Lee, B. (2016). "The mind behind crowdfunding: an empirical study of speech emotion in fundraising success," in *Proceedings 37th International Conference on Information Systems (ICIS)* (Bangkok), 1–11.

Kimani, E., Murali, P., Shamekhi, A., Parmar, D., Munikoti, S., and Bickmore, T. (2020). "Multimodal assessment of oral presentations using hmms," in *ICMI '20* (New York, NY: ACM), 650–654.

Kisfalvi, V., and Pitcher, P. (2003). Doing what feels right: the influence of ceo character and emotions on top management team dynamics. *J. Manag. Inquiry* 12, 42–66. doi: 10.1177/1056492602250518

Koolagudi, S., and Rao, K. (2012). Emotion recognition from speech: a review. *Int. J. Speech Technol.* 15, 99–117. doi: 10.1007/s10772-011-9125-1

Landgraf, R. (2014). "Are you serious? irony and the perception of emphatic intensification," in *Proceedings of 4th International Symposium on Tonal Aspects of Languages (TAL 2014)* (Nijmegen), 91–94.

Lausen, A., and Hammerschmidt, K. (2020). Emotion recognition and confidence ratings predicted by vocal stimulus type and prosodic parameters. *Humanit Soc. Sci. Commun.* 7, 2. doi: 10.1057/s41599-020-0499-z

Liénard, J. S., and D. B. M. (1999). Effect of vocal effort on spectral properties of vowels. *J. Acoust. Soc. Am.* 106, 411–422. doi: 10.1121/1.428140

Liu, L., Suh, A., and Wagner, C. (2018). Empathy or perceived credibility? an empirical study on individual donation behavior in charitable crowdfunding. *Internet Res.* 28, 240. doi: 10.1108/IntR-06-2017-0240

Liu, X., and Xu, Y. (2014). "Body size projection by voice quality in emotional speechEvidence from Mandarin Chinese," in *Proceedings of 7th International Conference on Speech Prosody* (Dublin), 974–977. doi: 10.21437/SpeechProsody.2014-184

Ludusan, B., Wagner, P., and Włodarczak, M. (2021). Cue interaction in the perception of prosodic prominence: the role of voice quality. *Proc. Interspeech* 2021, 1006–1010. doi: 10.21437/Interspeech.2021-1357

Luengo, I., Navas, E., Hernáez, I., and Sánchez, J. (2005). Automatic emotion recognition using prosodic parameters. *Proc. Interspeech* 2005, 493–496. doi: 10.21437/Interspeech.2005-324

Lutzky, M., Schuller, G., Gayer, M., Krämer, U., and Wabnik, S. (2004). "A guideline to audio codec delay," in *Proceedings of 116th Audio Engineering Society Convention* (Berlin), 1–10.

Maran, T., Liegl, S., Moder, S., Kraus, S., and Furtner, M. (2021). Clothes make the leader! how leaders can use attire to impact followers' perceptions of charisma and approval. *J. Bus. Res.* 124, 86–99. doi: 10.1016/j.jbusres.2020.11.026

Maruschke, M., Jokisch, O., Meszaros, M., Trojahn, F., and Hoffmann, M. (2016). "Quality assessment of two fullband audio codecs supporting real-time

communication," in *Proceedings of the 18th International Conference on Speech and Computer SPECOM 2016* (Budapest), 571–579.

Ménard, L., Schwartz, J.-L., Bo,ë, L.-J., and Aubin, J. (2007). Articulatory-acoustic relationships during vocal tract growth for french vowels: analysis of real data and simulations with an articulatory model. *J. Phon.* 35, 1–19. doi: 10.1016/j.wocn.2006.01.003

Menezes, C., and Maekawa, K. (2006). "Paralinguistic effects on voice quality: a study in Japanese," in *Proceedings of the 3rd International Conference of Speech Prosody, Paper 049*, 1–4.

Mozziconacci, S. (2001). Modeling emotion and attitude in speech by means of perceptually based parameter values. *User Model Useradapt Interact.* 11, 297–326. doi: 10.1023/A:1011800417621

Niebuhr, O. (2021). "Advancing higher education practice by analyzing and training students vocal charisma: evidence from a danish field study," in *Proceedings of 7th International Conference on Higher Education Advances (Head 2021)* (Valencia), 1–8.

Niebuhr, O., and Michalsky, J. (2019). "Computer-generated speaker charisma and its effects on human actions in a car-navigation system experiment—or how steve jobs' tone of voice can take you anywhere," in *Procedings of Computational Science and Its Applications-ICCSA 2019, volume 11620 LNCS*, eds S. Misra, C. Torre, E. Tarantino, B. O. Apduhan, O. Gervasi, B. Murgante, E. Stankova, V. Korkhov, A. M. A. Rocha, and D. Taniar (Saint Petersburg: Springer Verlag), 375–390.

Niebuhr, O., Peters, B., Landgraf, R., and Schmidt, G. (2015). "The kiel corpora of "speech and emotion"—a summary," in *Tagungsband-DAGA 2015* (Rostock), 1011–1014.

Niebuhr, O., Reetz, H., Barnes, J., and Yu, A. C. L. (2020). *Fundamental Aspects in the Perception of f0*. Oxford: Oxford University Press.

Niebuhr, O., and Siegert, I. (2022). ""high on emotion "? how audio codecs interfere with the perceived charisma and emotional states of men and women," in *Elektronische Sprachsignalverarbeitung 2022. Tagungsband der 33. Konferenz, volume 103 of Studientexte zur Sprachkommunikation* (Sonderborg: TUD Press), 243–252.

Niebuhr, O., and Silber-Varod, V. (2021). How versatility performance influences perception of charismatic speech: a study on two israeli politicians. *Interact Stud.* 22, 303–342. doi: 10.1075/is.00007.nie

Niehof, J., and Niebuhr, O. (2022). "What's pitch got to do with it?—the connection between speech prosody and investor-pitch success," in *Proceedings of 1st International Conference on Tone and Intonation (TAI)* (Sonderborg), 288–292.

Parkins, R. (2012). "Gender and emotional expressiveness: an analysis of prosodic features in emotional expression," in *Griffith Working Papers in Pragmatics and Intercultural Communication, Vol. 5*, 46–54. Available online at: https://webarchive.nla.gov.au/awa/20130709003058/http://pandora.nla.gov.au/pan/85901/20130709-0954/Vol-5-Issue-1-complete_final.pdf

Pausewang Gelfer, M., and Young, S. R. (1997). Comparisons of intensity measures and their stability in male and female sneakers. *J. Voice* 11, 178–186. doi: 10.1016/S0892-1997(97)80076-8

Pentland, A. (2008). *Honest Signals: How They Shape our World*. Cambridge: MIT Press.

Pépiot, E. (2014). Male and female speech: a study of mean f0, f0 range, phonation type and speech rate in parisian french and american english speakers. *Speech Prosody* 7, 305–309. doi: 10.21437/SpeechProsody.2014-49

Pfitzinger, H. R., Amir, N., Mixdorff, H., and Bösel, J. (2011). "Cross-language perception of hebrew and german authentic emotional speech," in *Proceedings of 17th International Congress of Phonetic Sciences* (Hong Kong), 1586–1589.

Pfitzinger, H. R., and Kaernbach, C. (2008). Amplitude and amplitude variation of emotional speech. *Proc. Interspeech* 2008, 1036–1039. doi: 10.21437/Interspeech.2008-322

Pollak, P., and Behunek, M. (2011). "Accuracy of mp3 speech recognition under real-word conditions: experimental study," in *Proceedings of the International Conference on Signal Processing and Multimedia Applications* (Seville), 1–6.

Pratt, T. (2020). Embodying "tech": articulatory setting, phonetic variation, and social meaning. *J. Sociolinguist.* 24, 328–349. doi: 10.1111/josl.12369

Raju, V., V. V., Gurugubelli, K., Alluri, K. R., and Kumar Vuppala, A. (2018). "Differenced prosody features from normal and stressed regions for emotion recognition," in *2018 5th International Conference on Signal Processing and Integrated Networks (SPIN)* (Noida), 821–825. doi: 10.1109/SPIN42206.2018

Reddy, A., and Vijayarajan, V. (2020). Audio compression with multi-algorithm fusion and its impact in speech emotion recognition. *Int. J. Speech Technol.* 23, 277–285. doi: 10.1007/s10772-020-09689-9

Rendall, D., Kollias, S., Ney, C., and Lloyd, P. (2005). Pitch (f0) and formant profiles of human vowels and vowel-like baboon grunts: the role of vocalizer body size and voice-acoustic allometry. *J. Acoust. Soc. Amer.* 117, 944–955. doi: 10.1121/1.1848011

Rosenberg, A., and Hirschberg, J. (2009). Charisma perception from text and speech. *Speech Commun.* 51, 640–655. doi: 10.1016/j.specom.2008.11.001

Rozborski, B. (2007). "A preliminary study on the influence of sound data compression upon formant frequency distributions in vowels and their measurement," in *Proceedings of ICPhS XVI* (Saarbrücken), 1833–1836.

Scherer, K. R. (2013). Vocal markers of emotion: Comparing induction and acting elicitation. *Comput. Speech Lang.* 27, 40–58. doi: 10.1016/j.csl.2011.11.003

Schuller, B., Batliner, A., Steidl, S., and Seppi, D. (2011). Recognising realistic emotions and affect in speech: state of the art and lessons learnt from the first challenge. *Speech Commun.* 53, 1062–1087. doi: 10.1016/j.specom.2011.01.011

Schuller, B., Vlasenko, B., Eyben, F., Rigoll, G., and Wendemuth, A. (2009). "Acoustic emotion recognition: a benchmark comparison of performances," in *Proceedings of the IEEE ASRU-2009* (Moreno: IEEE), 552–557.

Siegert, I., Lotz, A. F., l. Duong, L., and Wendemuth, A. (2016). "Measuring the impact of audio compression on the spectral quality of speech data," in *Elektronische Sprachsignalverarbeitung 2016, volume 81 of Studientexte zur Sprachkommunikation* (Leipzig), 229–236.

Siegert, I., and Niebuhr, O. (2021a). Case report: women, be aware that your vocal charisma can dwindle in remote meetings. *Front. Commun.* 5, 135. doi: 10.3389/fcomm.2020.611555

Siegert, I., and Niebuhr, O. (2021b). "Speech signal compression deteriorates acoustic cues to perceived speaker charisma," in *Elektronische Sprachsignalverarbeitung 2021. Tagungsband der 32. Konferenz, volume 99 of Studientexte zur Sprachkommunikation* (Berlin: TUD Press), 1–10.

Simpson, A. P. (2009). Phonetic differences between male and female speech. *Lang. Linguist. Compass* 3, 621–640. doi: 10.1111/j.1749-818X.2009.00125.x

Steidl, S., Polzehl, T., Bunnell, H. T., Dou, Y., Muthukumar, P. K., Perry, D., et al. (2012). Emotion identification for evaluation of synthesized emotional speech. *Proc. Speech Prosody* 2012, 661–664. Available online at: https://www.isca-speech.org/archive/speechprosody_2012/steidl12_speechprosody.html

Strangert, E., and Gustafson, J. (2008). What makes a good speaker? subject ratings, acoustic measurements and perceptual evaluations. *Proc. Interspeech* 2008, 1688–1691. doi: 10.21437/Interspeech.2008-368

Sy, T., Horton, C., and Riggio, R. (2018). Charismatic leadership: Eliciting and channeling follower emotions. *Leadership Q.* 29, 58–69. doi: 10.1016/j.leaqua.2017.12.008

Tarasov, A., and Delany, S. J. (2011). "Benchmarking classification models for emotion recognition in natural speech: a multi-corporal study," in *Proceedings of the 9th IEEE FG* (Santa Barbara, CA), 841–846.

Terken, J., and Hermes, D. (2000). *The Perception of Prosodic Prominence*. Dordrecht: Springer Netherlands.

Touati, P. (1993). "Prosodic aspects of Political rhetoric," in *Proceedings of ESCA Workshop on Prosody* (Lund), 168–171.

Traunmüller, H., and Eriksson, A. (1995). *The frequency range of the voice fundamental in the speech of male and female adults*. Linguistics Department, University of Stockholm, Stockholm, Sweden.

Tur, B., Harstad, J., and Antonakis, J. (2021). Effect of charismatic signaling in social media settings: evidence from ted and twitter. *Leadership Q.* 33, 101476. doi: 10.1016/j.leaqua.2020.101476

Valin, J.-M., Maxwell, G., Terriberry, T. B., and Vos, K. (2013). "The opus codec," in *135th AES International Convention* (New York, NY).

Valin, J.-M., Vos, K., and Terriberry, T. (2012). *Definition of the opus audio codec*. RFC 6716.

van Son, R. J. J. H. (2005). A study of pitch, formant, and spectral estimation errors introduced by three lossy speech compression algorithms. *Acta Acustica United Acust.* 91, 771–778.

Vary, P., and Martin, R. (2006). *Digital Speech Transmission: Enhancement, Coding and Error Concealment*. Chichester: Wiley & Sons.

Ververidis, D., and Kotropoulos, C. (2006). Emotional speech recognition: resources, features, and methods. *Speech Commun.* 48, 1162–1181. doi: 10.1016/j.specom.2006.04.003

Waaramaa, T., Laukkanen, A., Airas, M., and Alku, P. (2010). Perception of emotional valences and activity levels from vowel segments of continuous speech. *J. Voice* 1, 4. doi: 10.1016/j.jvoice.2008.04.004

Wang, Y., Du, S., and Zhan, Y. (2008). "Adaptive and optimal classification of speech emotion recognition," in *2008 Fourth International Conference on Natural Computation, Vol. 5* (Washington: DC), 407–411. doi: 10.1109/ICNC.2008.995

Weiss, B., Burkhardt, F., and Geier, M. (2013). "Towards perceptual dimensions of speakers' voices: eliciting individual descriptions," in *Workshop on Affective Social Speech Signals (WASSS 2013)* (Grenoble).

Wörtwein, T., Chollet, M., Schauerte, B., Morency, L.-P., Stiefelhagen, R., and Scherer, S. (2015). "Multimodal public speaking performance assessment," in *Proceedings of the 2015 ACM on International Conference on Multimodal Interaction* (New York, NY), 43–50.

Xiph.Org Foundation (2014). *Speex: A Free Codec for Free Speech*. Available online at: https://gitlab.xiph.org/xiph/speex (accessed March 8, 2023).

Zhao, K., Zhou, L., and Zhao, X. (2022). Multi-modal emotion expression and online charity crowdfunding success. *Decis Support Syst.* 163, 113842. doi: 10.1016/j.dss.2022.113842

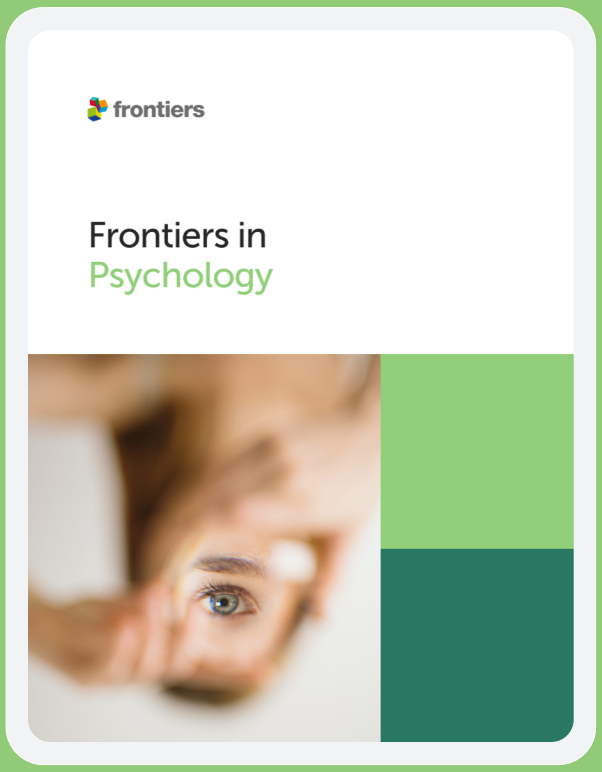
Frontiers in Psychology

Paving the way for a greater understanding of human behavior

The most cited journal in its field, exploring psychological sciences - from clinical research to cognitive science, from imaging studies to human factors, and from animal cognition to social psychology.

Discover the latest Research Topics

See more →



Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact