frontiers | Research Topics

Reimagining roles and identity in the era of human - AI collaboration

Edited by
Xi Chen, Ivan Wen, Qixing Qu and Wenjing Chen

Published in
Frontiers in Psychology

ISSN 1664-8714
ISBN 978-2-8325-7250-4
DOI 10.3389/978-2-8325-7250-4

Generative AI statement
Any alternative text (Alt text) provided alongside figures in the articles in this ebook has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

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

Reimagining roles and identity in the era of human - AI collaboration

Topic editors
Xi Chen — Yunnan University, China
Ivan Wen — University of Hawaii at Manoa, United States
Qixing Qu — University of International Business and Economics, China
Wenjing Chen — Beijing University of Posts and Telecommunications (BUPT), China

Citation
Chen, X., Wen, I., Qu, Q., Chen, W., eds. (2025). *Reimagining roles and identity in the era of human - AI collaboration*. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-8325-7250-4

Table of contents

TYPE Editorial
PUBLISHED 26 November 2025
DOI 10.3389/fpsyg.2025.1736730

OPEN ACCESS

EDITED AND REVIEWED BY
Rosanna E. Guadagno,
University of Oulu, Finland

*CORRESPONDENCE
Xi Chen
✉ 13610491@qq.com
Ivan Wen
✉ hsingwen@hawaii.edu
Qixing Qu
✉ qqxing@uibe.edu.cn
Wenjing Chen
✉ cwjbupt@bupt.edu.cn

RECEIVED 31 October 2025
REVISED 04 November 2025
ACCEPTED 18 November 2025
PUBLISHED 26 November 2025

CITATION
Chen X, Wen I, Qu Q and Chen W (2025) Editorial: Reimagining roles and identity in the era of human - AI collaboration. *Front. Psychol.* 16:1736730. doi: 10.3389/fpsyg.2025.1736730

Editorial: Reimagining roles and identity in the era of human - AI collaboration

Xi Chen[1]*, Ivan Wen[2]*, Qixing Qu[3]* and Wenjing Chen[4]*

[1]Management Science Department, School of Business and Tourism Management, Yunnan University, Kunming, China, [2]School of Travel Industry Management, Shidler College of Business, University of Hawai'i at Manoa, Honolulu, HI, United States, [3]School of International Trade and Economics, University of International Business and Economics, Beijing, China, [4]School of Economics and Management, Beijing University of Posts and Telecommunications, Beijing, China

KEYWORDS

roles, identity, human–AI collaboration, artificial intelligence, AI for social good

Editorial on the Research Topic
Reimagining roles and identity in the era of human - AI collaboration

Human civilization is entering an epoch of profound human–AI collaboration—an era in which interaction between humans and intelligent systems no longer belongs to speculation but defines a new frontier of interdisciplinary inquiry (Vaccaro et al., 2024). Within this emerging symbiosis, the essence of subjectivity and identity demands renewed scrutiny. The boundaries of the human—its agency, autonomy, and existential significance—are being redrawn in collaborative terrains where technology no longer serves merely as an instrument but participates intimately in cognition, perception, and decision-making (Fügener et al., 2022; Reinhardt et al., 2025).

Indeed, human–AI collaboration reshapes the very fabric of intersubjectivity. Artificial intelligence has evolved beyond a mechanical tool into a quasi-subjective partner in reasoning and creation (Hou et al., 2025). In this process, it unsettles established hierarchies of power, redistributes responsibility, and reconfigures the mechanisms of value co-creation (Wessel et al., 2025). Such transformation calls for reaffirming the distinctiveness of human cognition, emotion, and moral judgment—those fragile yet irreplaceable capacities that lend ethical texture to progress (Glickman and Sharot, 2025). At the same time, AI dissolves the once-stable boundary between reality and virtuality, transforming both the spaces and meanings of identity expression (Heinrich and Gerhart, 2025). Subjectivity now extends beyond the corporeal self into plural performances across digital dimensions—liberating yet perilous, emancipatory yet disciplinary. Hence, technological advancement must be tempered by humanistic care, preserving dignity within empowerment and conscience within innovation.

At its conceptual core, this rethinking of subjectivity invites deeper reflection on the nature of humanity and intelligence in the technological age. It is a dialogue that transcends disciplinary boundaries, drawing from sociology, psychology, management, communication studies, and computer science. Together, these fields seek to understand the co-evolution of human consciousness and artificial cognition. Three themes define this frontier: the psychological and interactive dynamics of human–AI collaboration; the repositioning of human uniqueness within intelligent ecosystems; and the ethical principles guiding digital identities in AI-mediated environments. In the end, these dimensions form the foundation for reimagining human subjectivity amid technological symbiosis.

Empirical research illuminates this landscape, revealing how personality, emotion, and resource dynamics shape human–AI relations. Liu and Chen find that Generation Z's chatbot-assisted purchases are shaped by extraversion, agreeableness, and openness, along with chatbot expertise and customization—underscoring the need for designs attuned to human individuality rather than uniform assumptions. Yu and Chang show that students' digital photograph hoarding arises from emotional attachment and fear of missing out, as AI tools increasingly serve as repositories of affect and memory. Han and Ren reveal that unequal access to AI can paradoxically enhance team productivity through complementary interaction, challenging the notion that equality in technology always yields optimal collaboration. Collectively, these studies expose the complex interplay of personality alignment, emotional mediation, and strategic asymmetry that transcends traditional human–human frameworks.

Beyond interpersonal dynamics, AI is also redrawing the contours of roles and agency in academic and professional life. Huang and Zhao demonstrate that AI literacy enhances wellbeing by fulfilling needs for autonomy, competence, and relatedness, thereby improving work–life balance and job satisfaction. Zhao and Huang extend this view, showing that AI literacy stimulates pedagogical innovation through strengthened attitudes, norms, and perceived control, moderated by resources and autonomy. Jiang et al. reveal that AI-resistant skills, network centrality, and proactive personality foster collaboration, while digital identity reconstruction reorganizes participation and authority. Together, these insights portray AI not merely as an instrument of efficiency but as a transformative agent that redefines human creativity and purpose.

Yet as AI permeates every stratum of life, it also exposes humanity to profound ethical and psychological dilemmas. Chen et al. propose governance models with adaptive trust-repair mechanisms—tailoring attribution and social support to failure contexts while using anthropomorphic cues to sustain resilience. Fu et al. call for frameworks that balance technological utility with emotional wellbeing, highlighting the fragility of end-of-life AI applications where algorithms intersect with grief and post-humous identity. Drawing on Foucauldian notions of subjectivation, they warn that AI mourning tools may reconstitute moral agency beyond death itself. Meanwhile, Thomas and Manalil underscore the urgency of algorithmic transparency to mitigate emotional coercion and cognitive dissonance. Their depiction of shadow banning as "digital silence" reveals its erosion of self-perception and autonomy, urging oversight of both visible and subtle algorithmic harms. Collectively, these perspectives affirm that effective governance must weave together trust, ethics, and psychological awareness to ensure that AI systems remain profoundly humane.

Taken together, these insights illuminate a profound reciprocity: humans endow artificial intelligence with creativity, purpose, and moral direction, even as AI amplifies human potential and reshapes the horizons of thought and collaboration. The evolving discourse on roles and identities thus offers forward-looking pathways for understanding how humanity constructs, safeguards, and enacts subjectivity within an increasingly algorithmic world. As intelligent systems weave themselves ever more deeply into the fabric of life, the imperative becomes clear—to ensure that innovation never eclipses emotion, conscience, and dignity (Bankins and Formosa, 2023). These reflections chart a transformative journey toward self-realization in the digital epoch and toward governance structures capable of reconciling technological power with ethical responsibility.

Ultimately, this corpus of scholarship converges upon the global aspiration for "AI for social good." It reminds us that the true horizon of progress does not reside in the perfection of machines, but in the deepening of our humanity—the enduring capacity to endow intelligence, whether human or artificial, with compassion, justice, and dignity.

Author contributions

XC: Writing – review & editing, Conceptualization, Writing – original draft. IW: Writing – original draft, Writing – review & editing. QQ: Writing – review & editing. WC: Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This work was supported by the National Education Science Planning Project-National General Project, "Rethinking and Reforming the Higher Education Evaluation Mechanism from the Perspective of Gender Social Imprinting" [grant number BSA250243].

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.

Generative AI statement

The author(s) declare that no Gen AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

References

Bankins, S., and Formosa, P. (2023). The ethical implications of artificial intelligence (AI) for meaningful work. *J. Business Ethics* 185, 725–740. doi: 10.1007/s10551-023-05339-7

Fügener, A., Grahl, J., Gupta, A., and Ketter, W. (2022). Cognitive challenges in human–artificial intelligence collaboration: investigating the path toward productive delegation. *Inform. Syst. Res.* 33, 678–696. doi: 10.1287/isre.2021.1079

Glickman, M., and Sharot, T. (2025). How human–AI feedback loops alter human perceptual, emotional and social judgements. *Nat. Hum. Behav.* 9, 345–359. doi: 10.1038/s41562-024-02077-2

Heinrich, M., and Gerhart, N. (2025). Privacy? What's that? Differences in privacy boundaries. *Online Inform. Rev*. 49, 97–115. doi: 10.1108/OIR-09-2023-0447

Hou, J., Wang, L., Wang, G., Wang, H. J., and Yang, S. (2025). The double-edged roles of generative ai in the creative process: experiments on design work. *Inform. Syst. Res*. doi: 10.1287/isre.2024.0937

Reinhardt, A., Matthes, J., Bojic, L., Maindal, H. T., Paraschiv, C., and Ryom, K. (2025). Help me, Doctor AI? A cross-national experiment on the effects of disease threat and stigma on AI health information-seeking intentions. *Comput. Hum. Behav.* 172:108718. doi: 10.1016/j.chb.2025.108718

Vaccaro, M., Almaatouq, A., and Malone, T. (2024). When combinations of humans and AI are useful: a systematic review and meta-analysis. *Nat. Hum. Behav*. 8, 2293–2303. doi: 10.1038/s41562-024-02024-1

Wessel, M., Adam, M., Benlian, A., Majchrzak, A., and Thies, F. (2025). Generative AI and its transformative value for digital platforms. *J. Manag. Inform. Syst.* 42, 346–369. doi: 10.1080/07421222.2025.2487315

frontiers | Frontiers in Digital Health

TYPE Original Research
PUBLISHED 07 July 2025
DOI 10.3389/fdgth.2025.1618169

Check for updates

OPEN ACCESS

EDITED BY
Xi Chen,
Yunnan University, China

REVIEWED BY
Hongying Du,
Kunming Medical University, China
Hugo Rodriguez Resendiz,
Autonomous University of Queretaro, Mexico

*CORRESPONDENCE
Kun Fu
kunfu@uic.edu.cn
Yuan Yuan
yuanyuan@uic.edu.cn

†These authors contributed equally to this work and share first authorship

RECEIVED 06 May 2025
ACCEPTED 17 June 2025
PUBLISHED 07 July 2025

CITATION
Fu K, Ye C, Wang Z, Wu M, Liu Z and Yuan Y (2025) Ethical dilemmas and the reconstruction of subjectivity in digital mourning in the age of AI: an empirical study on the acceptance intentions of bereaved family members of cancer patients.
Front. Digit. Health 7:1618169.
doi: 10.3389/fdgth.2025.1618169

Ethical dilemmas and the reconstruction of subjectivity in digital mourning in the age of AI: an empirical study on the acceptance intentions of bereaved family members of cancer patients

Kun Fu[1*†], Chenxi Ye[1†], Zeyu Wang[1†], Miaohui Wu[2], Zhen Liu[3] and Yuan Yuan[1*]

[1]School of Culture and Creativity, Beijing Normal-Hong Kong Baptist University, Zhuhai, Guangdong, China, [2]Faculty of Philosophy, Zhengzhou University, Zhengzhou, Henan, China, [3]School of Digital Art and Design, Chengdu Neusoft University, Chengdu, Sichuan, China

Introduction: With the rapid advancement of AI replication, virtual memorials, and affective computing technologies, digital mourning has emerged as a prevalent mode of psychological reconstruction for families coping with the loss of terminally ill patients. For family members of cancer patients, who often shoulder prolonged caregiving and complex ethical decisions, this process entails not only emotional trauma but also profound ethical dilemmas.
Methods: This study adopts the Unified Theory of Acceptance and Use of Technology (UTAUT) as its analytical framework, further integrating Foucauldian subjectivation theory and emotional-cognitive models. A structural path model was constructed to examine how ethical identification and grief perception influence the acceptance of AI-based digital mourning technologies. A total of 129 valid survey responses were collected and analyzed using Partial Least Squares Structural Equation Modeling (PLS-SEM).
Results: The findings indicate that performance expectancy, effort expectancy, social influence, and ethical concern significantly predict users' intention to adopt digital mourning technologies. Additionally, grief perception not only influences adoption intention but also directly affects actual usage behavior.
Discussion: This study highlights that the acceptance of AI-based digital mourning technologies extends beyond instrumental rationality. It is shaped by the interplay of emotional vulnerability and moral tension. The results contribute to a deeper understanding of the ethical and psychological dimensions of posthumous AI applications and provide valuable insights for future human-AI interaction design, digital commemoration systems, and the governance of end-of-life technologies.

KEYWORDS

AI digital mourning, UTAUT model, ethical concerns in AI, perception of grief, emprical study

1 Introduction

Digital mourning, as an emerging application of AI technology in end-of-life care, has gained traction as a form of commemorative practice following the death of cancer patients. This phenomenon encompasses a variety of technological forms—including AI-based digital replication (1), virtual reality (VR) memorial spaces, immersive interaction (2), and chatbots (3)—allowing bereaved family members to engage with the "digital identities" of deceased individuals within virtual environments. These technologies not only redefine traditional experiences of death but also reconstruct the cultural and psychosocial landscape of mourning itself (4).

While these AI products can simulate the deceased's behaviors and responses based on personal data, constructing a "ghost"-like digital mourning form through inference and prediction—thereby introducing novel support for emotional continuity and adaptive grief coping (5), they simultaneously generate a series of ethical tensions and psychosocial risks. Notably, algorithmic simulations of the deceased blur the ontological boundaries between life and death (6), potentially causing cognitive disconnection from mortality among the bereaved.Furthermore, AI-mediated mourning may foster a commercialized "affective outsourcing" (7)—where mourners' subjectivity becomes increasingly co-constituted, even subordinated to mechanical processes of memory management and emotional regulation. These developments compel a reexamination of two fundamental questions: What constitutes authentic grief? And to what extent can mourning—once a private, human-centered process—be technologized without compromising its existential significance and moral core?

In terms of form, digital mourning technologies provide more diverse avenues for memorialization, particularly under the integration of AI and virtual reality, where their roles in emotional companionship and memory reconstruction have gained increasing attention. However, for family members of cancer patients—who often endure prolonged caregiving and emotional exhaustion—this process may not signify healing; rather, it may exacerbate both ethical dilemmas and grief perception.

Cancer typically entails a slow and irreversible process of bodily deterioration, often accompanied by intense pain, a sense of medical futility, and the erosion of personal dignity (8). Family members, in such contexts, frequently undertake multiple roles: as emotional companions, caregiving executors, and ethical proxies in medical decision-making (9). The emotional burdens accumulated during this period rarely dissipate after the patient's death; instead, they often manifest in highly complex grief experiences—such as prolonged sadness, guilt, moral distress, or even post-traumatic symptoms.

Against this backdrop, the introduction of digital mourning technologies—such as AI-based replication and VR memorials—though envisioned as tools for emotional connection and memory continuity, may present unique ethical and psychological challenges for cancer-bereaved families. On one hand, digital identities are typically constructed from limited pre-death data and are prone to distortion or recomposition during algorithmic generation (10). The inconsistencies between replicated personas and real memories may create identity dissonance and a rupture in the sense of authenticity (11). On the other hand, for those whose emotional wounds from caregiving remain unhealed, the AI-mediated reproduction of the deceased's voice, image, or interactive behavior—while seemingly offering comfort (12)—may inadvertently trigger emotional flooding, grief recurrence, or even psychological retraumatization (13).

Moreover, cancer care often involves highly moralized decisions such as "when to let go" or "whether to prolong life," making the technical reconstitution of the deceased a potential catalyst for renewed existential reflection—Has death truly occurred? Has mourning reached completion? These questions evoke deeply entangled experiences of ethical unease (14) and grief perception (15).

Therefore, for bereaved family members of cancer patients, digital mourning is not merely a matter of behavioral adoption of new technologies. Rather, it constitutes a psychosocial mechanism at the intersection of ethical judgment, emotional processing, and technological identity. This constitutes the theoretical starting point of the present study.

While existing literature has primarily focused on the emotional and technical feasibility of such technologies, there remains a critical lack of analysis on how bereaved families conceptualize the interrelation between technology, ethics, and grief. In particular, the mechanisms through which grief experience interacts with ethical tensions in digital mourning have yet to be systematically theorized. The relationship between digital technologies and moral norms is complex and mutually constitutive. Technologies not only shape values and environments but are themselves embedded in and shaped by normative frameworks—a core focus of ethical analysis (16, 17).

To address these gaps, the present study constructs a technology acceptance model for bereaved family members based on the Unified Theory of Acceptance and Use of Technology (UTAUT). It incorporates Foucault's theory of subjectivation and phenomenological-ethical inquiry to critically frame the psychological and normative dimensions of digital mourning. By introducing ethical conflict perception and grief perception (ICG) as independent variables, this study seeks to empirically examine the extent to which AI-mediated mourning is accepted by bereaved family members of cancer patients.

2 Literature review and research hypotheses

2.1 Unified theory of acceptance and use of technology (UTAUT)

The Unified Theory of Acceptance and Use of Technology (UTAUT) was introduced by Venkatesh and colleagues in 2003. The main goal of this model was to combine the strengths of various previous models related to technology acceptance. By doing this, UTAUT aimed to improve the ability to explain and predict why users accept and use technology, as well as how they behave when using it. UTAUT integrates eight earlier models, including the Technology Acceptance Model (TAM), Theory of Planned Behavior (TPB), and Innovation Diffusion Theory (IDT), among others. It establishes a core framework based on performance expectancy, effort expectancy, social influence, and facilitating conditions, while incorporating gender, age, experience, and voluntariness as moderating variables to account for differences in technology acceptance across demographic groups (18). Subsequently, numerous scholars have extended the UTAUT framework by integrating contextual factors, such as cultural influence (19, 20), perceived risk (21), trust (22), and users' emotional responses (23, 24).

Since its inception, UTAUT has been widely applied across a variety of domains due to its strong predictive capabilities, including education (25), healthcare (26), e-government (27), fintech (28), and mobile internet (29). To further enhance its predictive scope, Venkatesh, Thong and Xu (23) proposed UTAUT2 adding new constructs such as hedonic motivation, price value, and habit to better account for technology adoption in consumer contexts. Many scholars have since built upon the UTAUT framework by integrating aspects like cultural influences, perceived risk, trust, and users' emotional responses. This has led to the model's enrichment across various academic fields and cultural contexts. These advancements have substantially deepened the theoretical understanding of UTAUT and broadened its practical relevance.

In recent studies, the UTAUT has been increasingly employed to explore user acceptance of emerging digital technologies such as artificial intelligence (30) and virtual reality (31). However, our review of current studies indicates that existing applications of the model often overlooks the ethical and emotional dimensions of technology acceptance. To address this gap, this study proposes an innovative extension of the UTAUT framework, demonstrating that the model not only effectively captures rational acceptance behavior but can also be integrated with variables related to emotions, ethics, and perceived risks to uncover the deeper psychological drivers behind technology adoption. Moving forward, as technological progress becomes more intertwined with social and ethical concerns, the continued integration and development of the UTAUT model will remain highly valuable both in theory and in practice.

2.2 Ethical issues in digital mourning

With the rapid advancement of artificial intelligence technologies, digital mourning has emerged as a novel form of commemoration and has been increasingly integrated into practices of end-of-life care and funerary culture (12). For instance, through AI-based replication, virtual memorial spaces, and voice-interactive systems, bereaved families can engage in immersive interactions with so-called "deathbots" representing the deceased (32).Specifically, we now categorize ethical issues into four interrelated dimensions, each supported by recent scholarly literature:

Identity Authenticity: AI-generated simulations may misrepresent the deceased's moral character, personality, or social roles, leading to a distortion of memory (33). Consent Ambiguity: Most platforms lack mechanisms for pre-mortem consent regarding digital data usage, creating unresolved issues around authorization (34). Emotional Manipulation: Extended AI-mediated interactions may cultivate emotional dependency, intensifying grief instead of alleviating it (35). Posthumous Data Rights: The commodification of digital remains has triggered ownership disputes between bereaved families and commercial providers (34).

Drawing on Foucault's concepts of disciplinary power and subjectivation (36), these technologies—while ostensibly therapeutic (37)—can standardize and regulate grieving behaviors. This creates a form of "programmed grief," where personal mourning becomes shaped by algorithmic design. As a result, the mourner's agency is displaced by technologically scripted responses, diminishing autonomy and reducing mourning to a reactive process. In this context, digital mourning functions not simply as a commemorative tool, but as a subtle apparatus of governance within the digital surveillance environment (38).

While digital mourning offers new mediums for emotional expression and psychological comfort, it also raises a host of ethical concerns—particularly in the domains of data privacy, AI-based personhood simulation, and emotional manipulation (1). Furthermore, the right to individualized mourning (39) remains ill-defined, and empirical studies on these topics are still sparse (40). Consequently, measuring users' ethical awareness—particularly whether they perceive digital mourning as a potential overreach into sensitive posthumous data—can reflect the tension between technological trust and moral anxiety.

Beyond data privacy, a more contentious issue lies in the ethical legitimacy of reconstructing a deceased person's identity via AI (41). Some platforms train large language models capable of mimicking the deceased's speech patterns, behavioral preferences, and even generating personalized responses (42), leading to what may be described as "simulated personhood." While these AI systems are often branded with narratives of "continued existence," a fundamental ethical question persists: are these systems genuine extensions of the deceased, or merely algorithmic performers? This ambiguity poses risks of eroding posthumous dignity, potentially undermining the very notion of "honoring the dead" (41). Moreover, the illusion of real continuity may interfere with healthy grief processing: users may become emotionally attached to AI-generated surrogates, leading to delayed psychological detachment, emotional dependency, or identity confusion (32). Thus, while such systems simulate connection, they may disrupt the natural course of mourning and reshape individuals' perceptions of death itself (12).

Despite their therapeutic claims, digital mourning platforms may engage in subtle forms of emotional governance. Their design often includes automated prompts—like birthday reminders or holiday messages—embedded with therapeutic intent (43). Yet these algorithmic interventions shape users' grief trajectories, potentially overriding personal timelines (44)). This raises a critical question: are these features truly tailored to individual grieving needs, or do they reflect a broader tendency toward the technological standardization of mourning? If perceived as excessive or manipulative, these interactions may erode user trust and reduce the likelihood of technology adoption. Consequently, perceived ethical tension may emerge as a key determinant of behavioral intention—warranting its integration into extended UTAUT models.

2.3 Grief perception and bereavement experience

In the context of illness-related death—especially in cases of cancer, where the disease is protracted, the process of decline is gradual yet evident, and the caregiving burden is particularly heavy—the psychological responses associated with bereavement tend to be significantly more complex than those triggered by sudden death. Prior studies indicate that family members bereaved by cancer often experience elevated levels of psychological distress, including symptoms of depression and anxiety, which are closely linked to their perceived suffering of the patient during the end-of-life period (45, 46). These family members commonly experience a prolonged emotional process that includes diagnosis, treatment, decline, and ultimately, death. This journey is characterized by anticipatory grief (47), anxiety related to ethical decision-making (48), and self-sacrificing caregiving actions (49), all of which frequently develop into profound grief after the loss (50). This grief is not a simple feeling but rather a complex psychological condition involving sadness, denial, anxiety, loneliness, and guilt. Its strength and how long it lasts can go well beyond typical grieving patterns and may appear as complicated grief (51).

Complicated grief, also known as prolonged grief disorder or delayed mourning, has been strongly associated with major depressive disorder, post-traumatic stress disorder (PTSD), and significant difficulties in social interactions (52). In some instances, it can worsen PTSD symptoms (53). This condition is frequently marked by an inability to let go of the deceased, denial of the death, ongoing difficulties in managing emotions, and the breakdown of life goals and trust in others (54–56). As Lichtenthal and colleagues have pointed out, for those who cared for cancer patients, grief is not just an emotional response. It often stems from the loss of their identity as a caregiver, their sense of ethical control, and how they see themselves in relation to others—leading to a type of grief that disrupts their sense of self, is hard to express, and deeply unsettling (57).

In recent years, researchers have increasingly focused on how people's understanding and experience of grief affect their behavior (58–61). Instead of only seeing grief as a result, a growing amount of research now considers how individuals perceive grief—often measured using the Inventory of Complicated Grief (ICG)—as a cognitive and emotional factor that can influence whether they adopt new technologies, participate in social activities, and make decisions involving risk (62). Specifically, when it comes to technologies used at the end of life and AI tools for remembrance, a person's individual experience of suffering can significantly shape how they think and evaluate things, their moral judgments, and the choices they make. For instance, some studies have found that whether people are willing to accept AI technologies for mourning is closely linked to their emotional ability to cope and how they interpret grief. Those who feel emotionally resistant or have ethical doubts tend to be less willing to use these technologies (63, 64).

In the case of immersive digital mourning technologies, this psychological mechanism becomes especially complex. On one hand, these technologies can provide spaces for ongoing emotional connection and the preservation of memories, and are often viewed as ways to ease grief and strengthen the feeling of closeness with someone who has passed away (65). On the other hand, they might reawaken unresolved emotional pain, potentially trapping individuals in a repetitive cycle of technological mourning (66). In their assessment of VR-based grief interventions, Pizzoli et al. (2) discovered that individuals with high scores on the Inventory of Complicated Grief (ICG) were more likely to experience cognitive dissonance and a blurring of reality when interacting with AI-generated representations of the deceased. This "knowing it's artificial, but emotionally unable to let go" experience weakens the therapeutic efficacy of the technology (2). When such mechanisms intersect with AI-facilitated reanimations of the deceased, individuals may find themselves torn between the longing to reconnect and the emotional overload that compels rejection of the digital representation. These findings reinforce the view that grief perception is not a neutral background condition but a decisive antecedent variable in technology acceptance.

Accordingly, this study incorporates Perception of Complicated Grief as a key independent variable within the extended UTAUT model to predict bereaved cancer family members' willingness to adopt AI-based digital mourning technologies. Here, we use the Inventory of Complicated Grief (ICG) (67) scale to measure perception of complicated grief. This model refinement aligns with cognitive-emotional decision-making theory, which assigns functional roles to affective variables, and responds to the unique moral-emotional entanglements of the digital mourning context. By introducing this construct, the study aims to go beyond rationalist acceptance models to offer a more psychologically grounded understanding of how grief and death experiences shape technology adoption in ethically charged domains.

2.4 Research questions and hypotheses

Based on the preceding literature and theoretical integration, this study aims to address the following four core research questions:

RQ1: Can the four core predictors in the original UTAUT model—performance expectancy (PE), effort expectancy (EE), social influence (SI), and facilitating conditions (FC)—effectively predict the behavioral intention (BI) and use behavior (UB) of bereaved family members of cancer patients toward AI-based digital mourning technologies?

RQ2: Building on the UTAUT model, do context-specific variables such as ethical concern (EC) and grief perception (ICG) significantly enhance the model's explanatory power? In other words, do these extended constructs contribute a statistically and theoretically meaningful increment to the prediction of behavioral intention?

RQ3: Do demographic variables (e.g., age, gender) serve as moderators between key technology perception variables and behavioral intention? How do such moderating effects reveal differentiated behavioral pathways among users facing emotionally intensive technologies?

RQ4: Is behavioral intention (BI) still the strongest predictor of actual use behavior (UB) in the context of AI commemorative systems? In other words, once users form an intention to use the technology, does it consistently translate into actual engagement?

To conduct empirical tests on these issues, the following research hypotheses are proposed. The corresponding diagrams are shown in Figure 1:

H1. Performance expectancy (PE) has a significant positive effect on behavioral intention (BI).

H2a: Effort expectancy (EE) has a significant positive effect on behavioral intention (BI).

H2b: Gender (GDR) negatively moderates the relationship between effort expectancy (EE) and behavioral intention (BI).

H2c: Voluntariness of use (Vuse) positively moderates the relationship between effort expectancy (EE) and behavioral intention (BI).

H3a: Social influence (SI) has a significant positive effect on behavioral intention (BI).

H3b: Voluntariness of use (Vuse) negatively moderates the relationship between social influence (SI) and behavioral intention (BI).

H4. Facilitating Conditions: Facilitating conditions (FC) have a significant positive effect on behavior intention (BI).

H5a: Ethical concern (EC) has a significant negative effect on behavioral intention (BI).

H5b: Age negatively moderates the relationship between ethical concern (EC) and behavioral intention (BI).

H6a: Grief perception (ICG) has a significant positive effect on behavioral intention (BI).

H6b: Grief perception (ICG) has a significant positive effect on use behavior (UB).

H6c: Gender (GDR) negatively moderates the relationship between grief perception (ICG) and behavioral intention (BI).

H7: Behavioral intention (BI) has a significant positive effect on use behavior (UB).

3 Research method

3.1 Survey method

In the early stage of questionnaire design, the research team organized a small expert consultation meeting and invited two front-line practice experts from Chongqing Medical University to participate and provide guidance. Based on clinical experience, experts have put forward targeted suggestions on issues such as the emotional responses of family members of cancer patients during the mourning process, their acceptance of technology, and possible ethical problems, and have improved the specific expression of the questionnaire. Make it more acceptable for family members. Based on the four core variables, this study added ethical care perception and pain perception as supplementary variables. The average well completion time is approximately 20 minutes. The data collection lasted for one

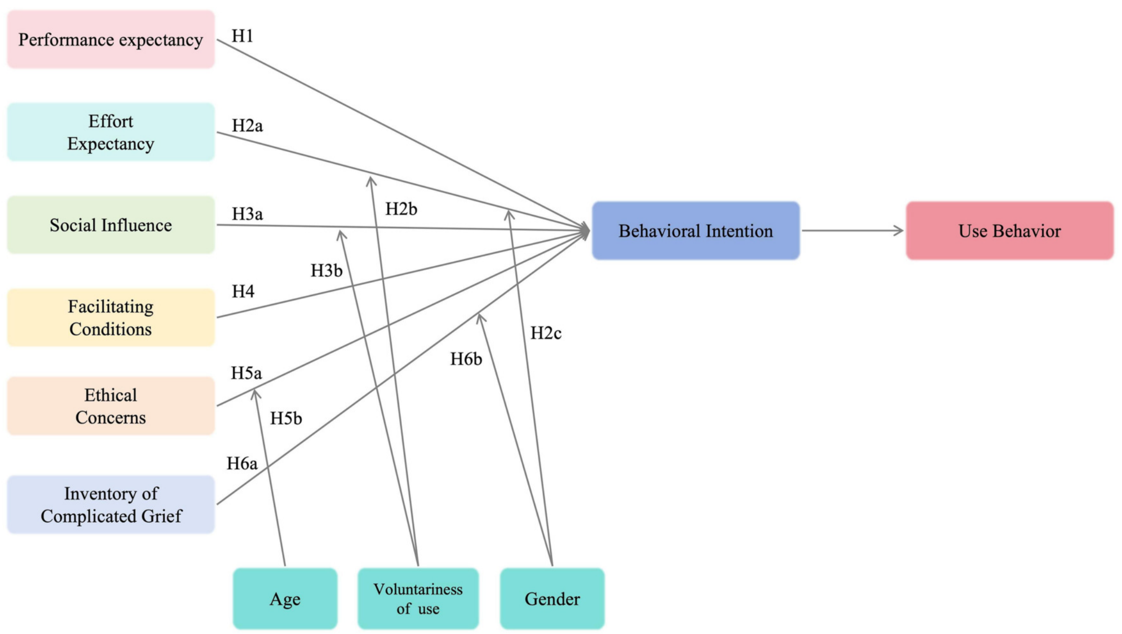


FIGURE 1
A proposed UTAUT model of AI acceptance among family members of deceased cancer patients.

week and a total of 137 responses were obtained. Among them, 129 were considered valid after data screening ($n = 129$).

3.2 Variable measurement

This study integrates the four core constructs of the Unified Theory of Acceptance and Use of Technology (UTAUT)—performance expectancy (PE), effort expectancy (EE), social influence (SI), and facilitating conditions (FC)—along with two original moderating variables (age and gender) and two additional context-specific variables: ethical concern (EC) and grief perception, measured via the Inventory of Complicated Grief (ICG). These constructs were adapted to reflect the psychological characteristics of bereaved family members of terminally ill patients. In total, six latent variables were measured.

Measurement items were developed by referencing and modifying the subdimensions of the original UTAUT scale proposed by Venkatesh et al., tailored to the specific context of bereavement and digital mourning. All constructs were measured using a five-point Likert scale, with 2–4 items per construct.

Participants (bereaved family members) were required to respond to all mandatory items. Example items included:"I believe AI-based mourning technologies can help me better commemorate my deceased loved one" (Performance Expectancy),"I find using AI mourning technologies difficult" (Effort Expectancy),"I think professionals (such as doctors or counselors) would recommend the use of AI mourning technologies" (Social Influence),"I can easily access guidance and assistance on how to use AI mourning technologies" (Facilitating Conditions).

3.3 Data analysis

To systematically explore the acceptance mechanisms of AI-based digital mourning among bereaved family members of cancer patients, this study employed a Partial Least Squares Structural Equation Modeling (PLS-SEM) approach. Data were analyzed using SmartPLS 27, which is well-suited for modeling complex path structures involving small samples, non-normal data, and moderated relationships.

Given the sensitivity of the study population—bereaved individuals with typically low public engagement, potential trauma triggers related to AI commemoration, ethical concerns, and limited technological exposure (1, 32)—the use of PLS-SEM is especially appropriate. The final dataset included 129 valid responses, and the Shapiro–Wilk test confirmed non-normal distribution for 43 out of 43 measurement items ($p < 0.05$). These conditions ($N < 200$ and significant non-normality) strongly justify the methodological fit of PLS-SEM, which remains robust under such constraints and does not rely on the assumption of multivariate normality.

This study first evaluated convergent validity by examining the factor loadings and average variance extracted (AVE) for each latent construct, and assessed internal consistency using composite reliability (CR). Subsequently, discriminant validity was tested using the Fornell–Larcker criterion to ensure adequate separation among the latent variables.

A structural model path diagram was generated, and the bootstrapping method was employed to assess key structural characteristics, including collinearity diagnostics, explanatory power (R^2), model fit (SRMR), and predictive relevance (Q^2). Finally, the significance of each hypothesized path in the extended UTAUT model was evaluated, which enabled the identification of significant relationships among the latent constructs and provided insights into the overall structural mechanism.

4 Digital research

4.1 Descriptive statistics of the sample

Following ethical approval from the Human Research Ethics Committee for Non-Clinical Faculties of Chengdu Neusoft University [Approval No. (CNU20241120)] and compliance with China's Personal Information Protection Law and institutional data governance standards, we administered the survey questionnaire distribution. A total of 207 questionnaires were distributed and collected in southern China. After manual data cleaning to remove invalid responses, 129 valid samples were retained, resulting in a valid response rate of 62.32%. Descriptive statistics of the sample were generated using SmartPLS.

Among the 129 valid respondents, 68 were female (52.71%) and 61 were male (47.29%). In terms of age distribution, the 18–25 age group constituted the majority of participants (62.02%), followed by the 26–30 age group (20.16%). Respondents aged over 50 years accounted for only 2.33% of the total sample. Demographic characteristics of the sample are summarized in Table 1.

4.2 Measurement model: reliability and validity assessment

This study employed the Partial Least Squares Algorithm function in SmartPLS 3.27 to evaluate the reliability and validity

TABLE 1 Descriptive statistics of respondent demographics.

Variable	Options	*n*	%
Gender	Male	61	47.29%
	Female	68	52.71%
Age	Lower 18	5	3.88%
	18–25	80	62.02%
	26–30	26	20.16%
	31–40	11	8.53%
	41–50	4	3.10%
	51–60	3	2.33%
	More than 60	0	0.00%
Education	Junior high school or below	1	0.78%
	Senior high school/Vocational school	9	6.98%
	University/Bachelor's degree	57	44.18%
	Postgraduate degree or above	62	48.06%

of each latent construct. Specifically, the analysis examined Cronbach's Alpha (CA), Composite Reliability (CR), and Factor Loadings for all items.

Validity assessment was conducted from two perspectives: convergent validity and discriminant validity. For convergent validity, the Average Variance Extracted (AVE) was calculated for each construct to assess the extent to which items reflect the intended latent variable. Discriminant validity was evaluated by comparing the square root of each construct's AVE with its correlations with other constructs, in accordance with the Fornell–Larcker criterion (68). Convergent validity results are detailed in Table 2, and Discriminant Validity results are presented in the Table 3.

The measurement model demonstrated satisfactory reliability, convergent validity, and discriminant validity through rigorous statistical validation. All constructs exhibited strong internal consistency (Cronbach's $\alpha > 0.70$) and convergent validity (AVE >0.50), aligning with thresholds defined by Hair et al. (68). Discriminant validity was confirmed through established criteria (e.g., HTMT ratios <0.85), ensuring distinctness among latent variables.

Moreover, the square roots of the AVE values for each construct were greater than their correlations with other constructs, and all factor loadings were higher than their respective cross-loadings—thus fulfilling the Fornell–Larcker criterion for discriminant validity (69).

The measurement model aligns with established psychometric standards for reliability, convergent validity, and discriminant validity, ensuring rigorous methodological grounding for the structural model's evaluation.

4.3 Structural model evaluation

After validating the measurement model, the study proceeded to examine the structural model, focusing on the model's predictive power and the causal relationships among latent constructs. The structural model was tested using SmartPLS 3.27, employing the bootstrapping procedure. The evaluation process included the following four steps:

(1) Collinearity assessment: Variance Inflation Factor (VIF) values were calculated to evaluate multicollinearity and the model's structural stability.
(2) Explanatory power: The Coefficient of Determination (R^2) was used to assess how well the exogenous constructs explained the variance in the endogenous variables.

TABLE 2 Convergent validity indicators for latent constructs (factor loadings, AVE, CR, cronbach's alpha).

Variables	Specific question items	Factor loading	Cronbach's alpha (CA)	Composite reliability (CR)	AVE
Behavioral intention (BI)	BI1	0.907	0.833	0.901	0.752
	BI2	0.877			
	BI3	0.815			
Ethical concern (EC)	EC1	0.905	0.884	0.928	0.811
	EC2	0.906			
	EC3	0.892			
Effort expectancy (EE)	EE1	0.875	0.839	0.919	0.850
	EE2	0.967			
Facilitating conditions (FC)	FC2	0.930	0.948	0.966	0.906
	FC3	0.963			
	FC4	0.962			
Performance expectancy (PE)	PE1	0.939	0.944	0.964	0.899
	PE2	0.959			
	PE3	0.946			
Social influence (SI)	SI1	0.911	0.781	0.901	0.820
	SI4	0.900			

TABLE 3 Discriminant validity assessment based on Fornell–Larcker criterion.

Construct	AGE	BI	ICG	EC	EE	FC	GDR	PE	SI	UB	Vuse
AGE	1										
BI	0.099	0.867									
ICG	0.05	0.433	1								
EC	−0.041	0.701	0.363	0.901							
EE	−0.078	−0.167	−0.41	−0.177	0.922						
FC	−0.08	0.399	0.413	0.387	−0.117	0.952					
GDR	−0.132	−0.004	0.103	0.051	−0.049	0.106	1				
PE	0.154	0.655	0.351	0.562	−0.26	0.2	−0.047	0.948			
SI	−0.019	0.593	0.271	0.548	−0.199	0.229	−0.023	0.464	0.906		
UB	0.067	0.758	0.447	0.607	−0.239	0.208	0.015	0.673	0.466	1	
Vuse	0.02	0.775	0.375	0.69	−0.188	0.339	−0.09	0.658	0.532	0.774	1

(3) Model fit: The Standardized Root Mean Square Residual (SRMR) was calculated as an index of model fit.
(4) Predictive relevance: The Construct Cross-Validated Redundancy (Q^2) was computed to evaluate the predictive relevance of the structural model (70).

These four indicators jointly assess the adequacy, explanatory power, and predictive performance of the model. In addition, the analysis of path coefficients, as well as direct and indirect effect sizes, was conducted to further evaluate the relationships among latent constructs. This step enables the study to address the research questions, test the proposed hypotheses, and determine the relative contribution of each independent variable to the acceptance of AI-based mourning technologies among bereaved family members.

According to the PLS-SEM framework, the model includes the following variables:

Exogenous latent constructs: Performance Expectancy (PE), Effort Expectancy (EE), Social Influence (SI), and Facilitating Conditions (FC).

Endogenous latent constructs: Behavioral Intention (BI) and Use Behavior (UB).

Observed moderating variables: Age, Gender, and Voluntariness of Use (Vuse).

Together, these components form the structural model used to explain and predict acceptance behavior toward AI-driven digital mourning technologies among family members of deceased cancer patients.

4.3.1 Collinearity diagnostics

In Partial Least Squares (PLS) data analysis, the Variance Inflation Factor (VIF) serves as a critical indicator for assessing potential multicollinearity within the structural model. As defined by Hair et al. in the context of SmartPLS-based modeling, a VIF value of 5 or higher indicates serious multicollinearity, whereas a VIF value of 3 or higher may suggest potential multicollinearity concerns that warrant further scrutiny (71).

As shown in the Table 4, all VIF values for the latent constructs in the model are below the threshold of 5, indicating that there is no severe multicollinearity among the variables. This finding validates the rationality of the questionnaire design, particularly the construct-specific item development strategy. Moreover, it suggests that the questionnaire items effectively differentiate between distinct latent dimensions, thereby minimizing the risk of estimation bias or model distortion caused by collinearity.

4.3.2 Evaluation of explanatory power

PLS-SEM employs ordinary least squares (OLS) regression to estimate path coefficients and factor loadings, aiming to maximize the explained variance (R^2) of endogenous constructs. This approach is particularly suitable for complex models and small samples, effectively capturing causal relationships among latent variables. According to Hair et al., the explanatory power of structural models can be categorized into three levels: $R^2 \geq 0.75$ (substantial), 0.50 (moderate), and 0.25 (weak) (72).

TABLE 4 Collinearity statistics of the structural model (VIF).

Specific items	VIF
EC1	2.376
EC2	2.613
EC3	2.548
EE22	2.093
EE33	2.093
FC1	2.826
FC2	3.505
FC3	3.947
GDR	1.000
PE1	3.237
PE3	3.473
PE4	2.284
SI1	1.695
SI4	1.695
UB	1.000
Vuse	1.000

TABLE 5 Coefficient of determination (R^2).

Endogenous variable	R-square	R-square adjusted
BI	0.770	0.745
UB	0.614	0.605

TABLE 6 Standardized root mean square residual (SRMR).

Model type	Saturated model	Estimated model
SRMR	0.077	0.078

As shown in the Table 5, the R^2 value for Behavioral Intention (BI) is 0.770, indicating that exogenous variables such as performance expectancy and effort expectancy collectively explain 77.0% of the variance in BI. This exceeds the typical explanatory power observed in conventional UTAUT applications, which usually ranges between 50% and 60%. The adjusted R^2 value of 0.745 further confirms the model's explanatory strength even after accounting for degrees of freedom, suggesting that the model is robust with respect to both variable count and sample size.

Similarly, the R^2 value for Use Behavior (UB) is 0.614, with an adjusted R^2 of 0.605. This indicates that the model explains 60.5% of the variance in actual use behavior, reflecting a relatively high level of explanatory power even after considering the interrelationships among the variables.

Taken together, these results demonstrate that the model possesses strong predictive capacity for the endogenous variables, supporting its validity for explaining user acceptance of AI-based applications in emotionally complex domains such as digital mourning.

4.3.3 Model fit evaluation

This study adopted the Standardized Root Mean Square Residual (SRMR) to assess the overall model fit. According to the criteria proposed by Henseler and Sarstedt, an SRMR value below 0.14 indicates acceptable model fit. The SRMR value of 0.078 (Table 6) indicates good model fit (73).

4.3.4 Predictive relevance (Q^2) evaluation

Predictive relevance (Q^2) is a key indicator in PLS-SEM used to assess the model's predictive validity. The Q^2 value ranges from negative infinity to 1, with higher values indicating stronger predictive relevance. In this study, the PLSpredict procedure was applied to compute Q^2 values. As shown in Table 7, the Q^2 values for the two endogenous latent variables were Behavioral Intention (BI) = 0.673 and Use Behavior (UB) = 0.613. Since both values are greater than zero, the results confirm that the exogenous constructs in the model exhibit adequate predictive relevance for the endogenous constructs.

5 Hypothesis testing results

This study applied bootstrapping with 5,000 resamples to estimate the path coefficients and assess their statistical significance within the structural model. The significance threshold was determined by T-statistics greater than 1.96 and p-values less than 0.10, with $p < 0.05$ being considered the standard for robust significance. The validity of each hypothesis was evaluated based on these criteria.

Additionally, the magnitude of each path coefficient indicates the relative strength of influence exerted by the independent variables on the dependent constructs. The results of hypothesis testing are summarized in Table 8, and the bar chart (Figure 2) summarizes the β coefficients and hypothesis testing results of all paths. Color coding is used to distinguish supported and unsupported hypotheses, as well as a simple slope interaction graph depicting the trajectories of behavioral intent (BI) under different independent variables (PE, EE, SI, FC, EC, ICG). It provides an intuitive understanding of path strength and directionality (see Figure 3).

TABLE 7 Predictive relevance (Q^2) results for the structural model.

Endogenous variable	Q^2predict
BI	0.673
UB	0.613

5.1 Hypotheses and interpretations

Based on the extended UTAUT model integrating both ethical and emotional variables, this study proposed a total of 13 hypothesis paths, of which 11 were statistically supported. These findings confirm that both affective and ethical factors play a critical role in shaping the behavioral intentions of bereaved family members toward AI-based digital mourning technologies.

First, performance expectancy (PE) was found to have a significant positive effect on behavioral intention (BI) (H1, $\beta = 0.150$, t = 2.015, $p = 0.044$), consistent with Venkatesh et al. (18), who argue that users' beliefs about the utility of a technology directly influence their intention to adopt it. In the context of digital mourning, family members who believe that AI technologies can alleviate grief or help restore emotional bonds are more inclined to accept their use.

Effort expectancy (EE) also exhibited a significant positive effect on BI (H2a, $\beta = 0.219$, t = 2.494, $p = 0.013$), suggesting that under emotionally intense circumstances, such as bereavement, individuals tend to value the ease of use and low emotional burden of new technologies. This is aligned with prior findings that emphasize the emotional benefits of user-friendly systems.

Social influence (SI) showed a significant impact on BI (H3a, $\beta = 0.138$, t = 1.981, $p = 0.048$), indicating that decisions around AI-based mourning are influenced not only by personal beliefs but also by the opinions of family, friends, and healthcare professionals. Moreover, voluntariness of use (Vuse) significantly and negatively moderated the relationship between SI and BI (H3b, $\beta = -0.134$, t = 2.660, $p = 0.008$), revealing that first-time users rely more heavily on external opinions, whereas more experienced users tend to form more autonomous judgments—reflecting an increase in user independence with experience.

Interestingly, facilitating conditions (FC) were found to have a significant negative effect on actual behavior Intention (BI) (H4, $\beta = -0.168$, t = 2.241, $p = 0.025$). While this contradicts the traditional UTAUT model assumption that facilitating

TABLE 8 An extended UTAUT model of acceptance and Use of AI-based mourning technologies Among bereaved families of cancer patients.

Hypothesis	Paths	Path coefficient (β)	Sample mean (M)	Standard deviation (STDEV)	T statistics (|O/STDEV|)	P values	Hypothesis testing
H1	PE -> BI	0.150	0.141	0.075	2.015	0.044	Supported
H2a	EE -> BI	0.219	0.201	0.088	2.494	0.013	Supported
H2b	GDR x EE -> BI	−0.206	−0.200	0.114	1.802	0.072	Not supported
H2c	Vuse x EE -> BI	−0.066	−0.065	0.047	1.419	0.156	Not supported
H3a	SI -> BI	0.138	0.134	0.07	1.981	0.048	Supported
H3b	Vuse x SI -> BI	−0.134	−0.133	0.05	2.66	0.008	Supported
H4	FC -> BI	−0.168	−0.164	0.075	2.241	0.025	Supported
H5a	EC -> BI	−0.227	−0.226	0.08	3.386	0.001	Supported
H5b	AGE x EC -> BI	−0.108	−0.107	0.054	1.98	0.048	Supported
H6a	ICG -> BI	0.283	0.278	0.088	3.222	0.001	Supported
H6b	ICG -> UB	0.198	0.198	0.069	2.893	0.004	Supported
H6c	GDR x ICG -> BI	−0.235	−0.233	0.108	2.178	0.029	Supported
H7	BI -> UB	0.737	0.736	0.061	12.098	0.000	Supported

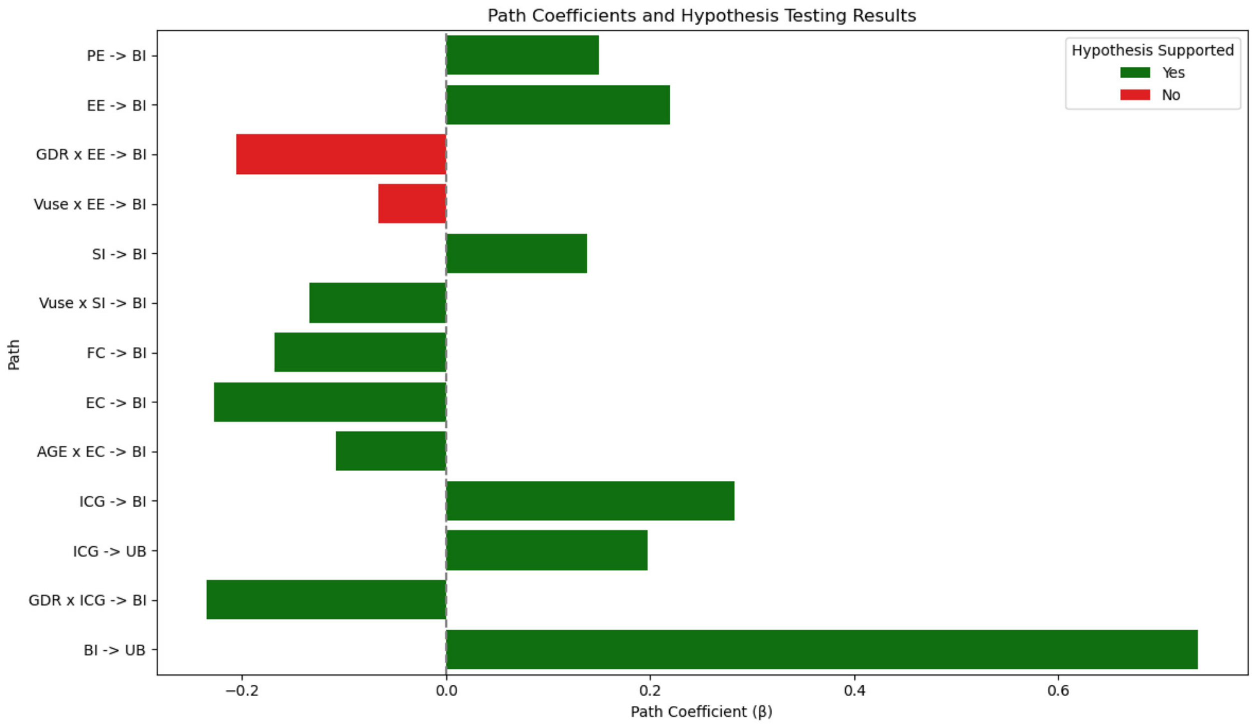


FIGURE 2
Path coefficients and hypothesis lesting results.

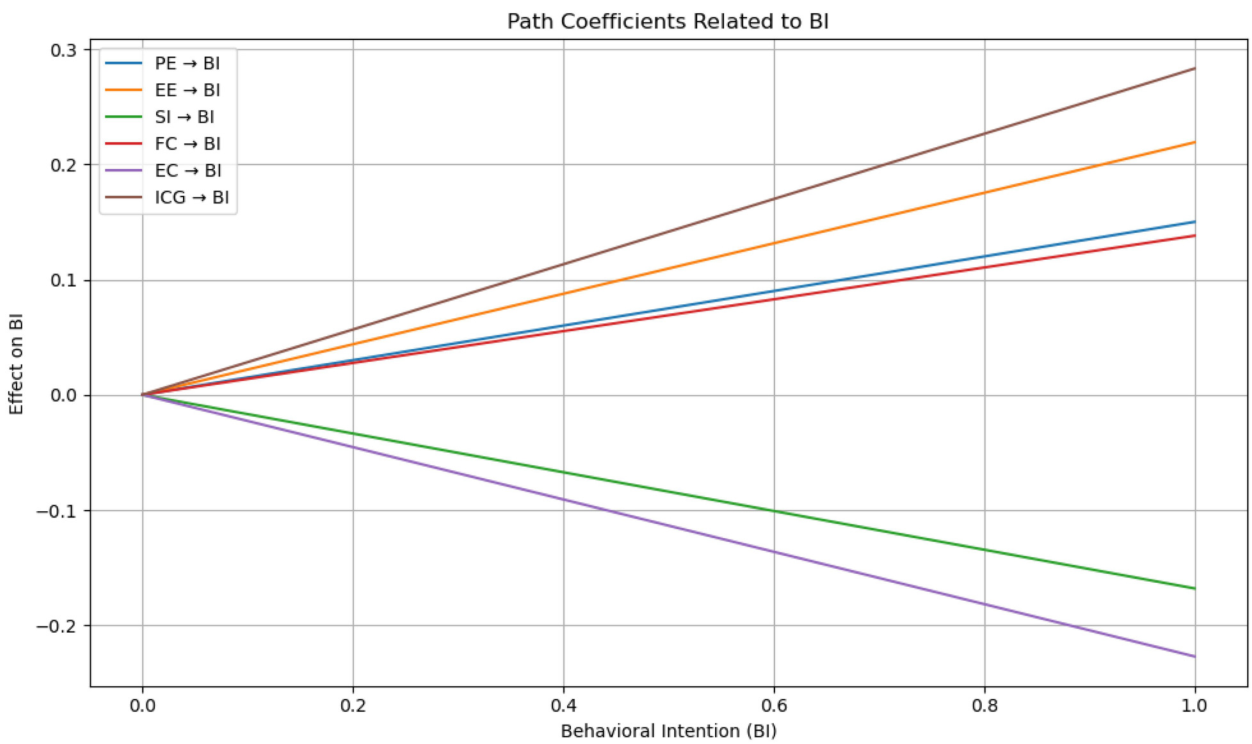


FIGURE 3
Dynamic path coefficients of behavioral intention (BI) determinants.

conditions promote behavioral adoption, it reveals a unique dynamic within the digital mourning context.

Ethical concern (EC) had a significant negative effect on behavioral intention (H5a, $\beta = -0.227$, $t = 3.386$, $p = 0.001$), echoing discussions in Chapter 2 that ethical considerations are central to digital mourning acceptance. Additionally, age was found to negatively moderate this relationship (H5b, $\beta = -0.108$, $t = 1.980$, $p = 0.048$), suggesting that older individuals may be more sensitive to ethical issues, thereby weakening the effect of ethical concern on their intention to adopt the technology.

Grief perception, as measured by the Inventory of Complicated Grief (ICG), significantly and positively influenced both behavioral intention (H6a, $\beta = 0.283$, $t = 3.222$, $p = 0.001$) and use behavior (H6b, $\beta = 0.198$, $t = 2.893$, $p = 0.004$). This supports the emotional activation hypothesis presented in Chapter 2—namely, that individuals experiencing higher levels of grief are more likely to engage with digital tools as a form of emotional compensation.

Furthermore, gender (GDR) negatively moderated the relationship between grief perception and behavioral intention (H6c, $\beta = -0.235$, $t = 2.178$, $p = 0.029$), indicating that gender-based psychological or emotional mechanisms may reduce the impact of grief perception on decision-making. Finally, behavioral intention strongly predicted actual use behavior (H7, $\beta = 0.737$, $t = 12.098$, $p < 0.001$), confirming the robust predictive power of intention in the context of AI-assisted mourning and supporting the structural validity of the UTAUT framework.

5.2 Unsupported hypotheses and interpretations

Despite most paths being statistically significant, two moderating hypotheses were not supported. Specifically, the moderating effects of gender and voluntariness of use on the relationship between effort expectancy and behavioral intention did not reach significance.

The first unsupported hypothesis was H2b, which posited a negative moderating effect of gender (GDR) on the relationship between effort expectancy (EE) and behavioral intention (BI) ($\beta = -0.206$, $t = 1.802$, $p = 0.072$). Although this value approached the significance threshold, it failed to meet the statistical cutoff. This suggests that in the context of digital mourning technologies for bereaved cancer families, perceptions of technological ease-of-use did not differ significantly across genders.

The second unsupported path, H2c, tested the moderating effect of voluntariness of use (Vuse) on the EE–BI relationship and was also not significant ($\beta = -0.066$, $t = 1.419$, $p = 0.156$). This implies that participants'prior experiences with similar technologies had no substantial influence on the relationship between their perceived ease of use and intention to adopt AI mourning tools.

These two unsupported hypotheses collectively reveal that effort expectancy, as a construct of instrumental reasoning, may be less susceptible to modulation by demographic or affective variables in emotionally intense contexts.

6 Discussion

6.1 Discussion on path assumptions

In this study, Hypothesis H1 is supported: Performance Expectancy (PE) exerts a significant positive effect on Behavioral Intention (BI), aligning with the original UTAUT model and indicating that users are more inclined to adopt AI-based digital mourning technologies when they believe such tools can effectively alleviate grief. This finding is consistent with Davis's (74) foundational insight that PE serves as a core driver of technology acceptance, often showing strong β correlations ranging from 0.63 to 0.85. However, the β value observed in this study falls below the typical range reported in UTAUT2, where Venkatesh et al. (23) noted that PE→BI path coefficients commonly exceed 0.3. This suggests that in the context of digital mourning, the perceived functional value of technology is subordinated to emotional needs, mirroring a similar attenuation trend observed in studies of medical AI (75).

For Hypothesis H2a, the positive impact of Effort Expectancy (EE) on BI reaffirms the foundational framework of UTAUT, suggesting that improvements in usability can directly enhance acceptance intention. This aligns with findings from the TAM2 extension, where EE typically influences BI indirectly via cognitive instrumental processes. However, both H2b and H2c, which test the moderating roles of gender and user experience on EE respectively, are not supported. This contradicts the original UTAUT model's conclusion that "gender moderates EE" (18). A plausible explanation lies in the emotional intensity of mourning behaviors, which may diminish individual differences, a pattern consistent with Li et al.'s (2023) findings in AI-mediated mental health contexts.

Hypothesis H3a, examining Social Influence (SI), is also supported, suggesting that normative pressure from friends, family, or society plays a facilitating role in the adoption of AI mourning technologies. Notably, H3b—which tests the interaction effect of user experience and SI on BI—is significant and negatively signed. This implies that more experienced users are less susceptible to social influence, which aligns with Venkatesh et al.'s (2003) moderation logic: experienced users tend to rely more on their autonomous judgment than on external cues.

Hypothesis H4 regarding Facilitating Conditions (FC) is supported, with a negative path coefficient indicating that environmental or resource-related obstacles (e.g., limited access to digital services) significantly reduce behavioral intention (BI). This reinforces the core UTAUT assumption that FC affects either BI directly or Use Behavior (UB) indirectly. However, the absolute β value is lower than that reported in some revised models. For instance, Dwivedi et al. (19) reported a path coefficient of approximately −0.34 for FC→BI. That indicates, usage of emotionally sensitive technologies, such as AI commemoration systems, may depend more on an individual's psychological readiness than on practical resources like access to devices or training. Even with available support, unresolved grief or ethical concerns can hinder actual use. Conversely, focusing heavily on the technical aspects of these systems might evoke negative emotional reactions or ethical objections, thereby reducing the likelihood of their adoption. These findings indicate that promoting the acceptance of these technologies requires attention to both practical support and users' emotional states, as well as ensuring that the technology aligns with their values.

Contrary to classical UTAUT findings (76), this study observed the disappearance of gender's moderating effect on the relationship between effort expectancy (EE) and behavioral intention (BI). This deviation may stem from the intense psychological distress inherent in cancer-related bereavement (45), which potentially overrides gender-specific behavioral patterns. Under such high-emotional-intensity conditions, both male and female bereaved individuals prioritize emotional security and existential authenticity over operational convenience, leading to a homogenization of technology evaluation criteria. This aligns with Suo et al.'s (2025) proposition that grief contexts neutralize gender disparities through an emotional homogenization effect.

Furthermore, voluntariness of use (Vuse) failed to moderate the EE→BI path—a finding resonant with Harbinja's ethical legitimacy threshold theory: "Users must first cross an ethical legitimacy threshold before evaluating usability in emotionally high-risk technologies" (77). This underscores that in digital mourning—a domain characterized by affective and ethical salience—utilitarian factors (e.g., ease of use) become secondary to existential concerns. The result corroborates Attuquayefio and Addo's (78) revised UTAUT framework, wherein moderating effects attenuate in high-stakes contexts. Digital mourning thus operates as an affective boundary condition, diminishing demographic sensitivity to functional attributes.

6.2 Principal findings

This study constructs an extended technology acceptance model for digital mourning within the UTAUT framework by incorporating two new variables: Perceived Grief (ICG) and Ethical Perception (EC). The empirical findings reveal a systematic transformation of traditional moderation mechanisms under high-sensitivity contexts. The theoretical contributions can be summarized in two key areas:

a. Reconfiguration of Acceptance Hierarchies Driven by Technology Sensitivity:

   Classic UTAUT theory posits that demographic variables such as gender, age, and user experience exert significant moderating effects on the core acceptance paths (18). However, our study finds that such traditional moderators lose explanatory power in emotionally sensitive contexts. Specifically, gender does not significantly moderate the path between Effort Expectancy (EE) and Behavioral Intention, while user experience negatively moderates the path from Social Influence (SI) to Behavioral Intention. This directly contradicts findings in consumer technology contexts, where experience tends to reinforce social conformity (23). This paradox can be interpreted through the lens of Technology Sensitivity Theory: when technologies intervene in emotionally charged scenarios (e.g., mourning, healthcare), users shift from a "function-first" to an "emotion-ethics-first" decision logic. As a result, demographic moderators become selectively operative only along emotion-ethical pathways, forming a context-dependent moderation filtering mechanism (75). Correspondingly, our findings show that age significantly strengthens the inhibitory effect of ethical perception, while gender attenuates the motivational effect of grief perception—indicating a reversal of traditional functional moderators. These findings challenge the universal applicability of UTAUT's moderation logic and propose new theoretical standards for researching high-sensitivity technologies.

b. The Emotional Authenticity Paradox and Ethical Intergenerational Effects in AI Mourning Technology Acceptance:

   This study also identifies two distinctive moderation effects absent from prior research: the emotional authenticity paradox and the ethical intergenerational effect. First, the negative moderation of social influence by usage experience (H3b) indicates that individuals with more digital mourning experience exhibit greater resistance to socially normative persuasion. This finding stands in sharp contrast to educational technology research, where increased experience tends to enhance social compliance (79). This divergence may stem from the inherently private nature of mourning: as users accumulate technological experience, they develop an awareness of emotional autonomy, becoming increasingly vigilant toward external interventions that might compromise the authenticity of their grief.

   Second, the study reveals a pronounced intergenerational ethical effect: age exerts a stronger negative moderation on ethical perception than on traditional predictors such as Effort Expectancy (typically $|\beta| < 0.05$). Older users tend to prioritize ethical boundaries over functional convenience in technology adoption decisions. This aligns with findings by Li et al. (80), who observed that "digital natives" focus more on usability, whereas "digital immigrants" emphasize ethical limits. These insights suggest the need to recalibrate UTAUT's moderation mechanism by incorporating an "ethical weighting coefficient" for age-related analyses in morally sensitive technological contexts.

6.3 Technical governance and suggestions

In terms of Chinese law, the data of the deceased is regarded as an object of property rights (Article 994 of the Civil Code), but the essence of digital mourning is to maintain the emotional connection between the living and the deceased. Therefore, the "maintaining connection" principle proposed by Chen Xiyi can be drawn upon to establish a "special management right for digital Remains" (81). The immediate family members of the deceased can be regarded as default managers to exercise data access rights in private mourning Spaces. When it comes to public mourning, a multi-party consultation committee should be established to balance personal emotions and public interests.

This mechanism can draw on the transitional arrangements of the European Union for deadbots (82), but it places more emphasis on the sustainability of the relationship rather than the disposal of the heritage.

At the social level, it is also very important to cultivate certain pre-social resilience. Incorporate the "empathy network" into the public crisis response system, such as opening digital mourning entrances after major accidents, or developing and advocating digital life education courses to guide young people to understand the boundaries of AI mourning technology first.

At the level of digital application, medical AI retains the "non-algorithmic" emotional space of doctor-patient interaction. Digital mental health tools should set protection thresholds for the mourning process to replace automated processes and avoid the formation of "cognitive dilemmas". An adaptive interface for the mourning stage can also be developed. Users' usage rights can be set to expand step by step based on the duration of use. First-time users cannot directly access all AI mourning services. The platform will proactively guide users to reach a moral consensus and improve the moral mechanism.

Furthermore, the research suggests that the deceased could sign an agreement during their lifetime to prohibit commercial or non-commercial digital revivals. For historical figures, certain ethical reviews are conducted through relevant experts and scholars.

7 Conclusion

7.1 Summary of key findings

This study used the UTAUT model to systematically investigate how bereaved family members accept and use AI-based digital mourning technologies. By adding ethical concerns and grief perception to the model and using PLS-SEM for data analysis, the research demonstrated that perceived usefulness, perceived ease of use, social influence, ethical considerations, and emotional distress significantly affect both the intention to use and the actual use of these technologies. The study also found that age, gender, and whether the use of the technology was voluntary or not, influence this acceptance in complex ways, highlighting the many factors that affect technology adoption in emotionally charged situations.

Going beyond these statistical results, the study uses Foucault's theories on how individuals become subjects to interpret digital mourning not just as a tool for coping with emotions, but also as a system that can shape behavior. AI commemoration technologies provide personalized ways to remember the deceased and offer emotional support, but they also subtly guide mourning into a digital practice that is structured by computational processes, interactions, and ongoing engagement. Consequently, the bereaved individual, who once expressed grief spontaneously, increasingly becomes a 'user' within a technological framework, with their mourning process and emotional pace influenced by the logic of these platforms. Digital mourning, therefore, serves not only as a source of comfort but also as a subtle mechanism of control.

7.2 Limitations and future work

In this study, the dominance of young participants (aged 18–30) inherently limited the ability of the research to capture intergenerational dynamics in mourning practices. The specific reason for this study is that the elderly often have deeper intergenerational traumatic memories, giving mourning behavior the significance of "family continuity", and they have a poor acceptance of the research questionnaire during the investigation period. Influenced by the trend of personalization, the youth group pays more attention to self-repair. Therefore, in the process of filling out the questionnaire, the proportion of the youth group is relatively large. This imbalance introduces a potential selection bias, favoring perspectives centered on individualistic coping and self-repair, which may not fully represent the communal or legacy-oriented mourning practices often observed among older adults. Future research must prioritize developing culturally sensitive and accessible methodologies (e.g., qualitative interviews, facilitated discussions, or alternative data collection formats) specifically designed to engage elderly populations and capture the richness of their grief experiences, particularly concerning intergenerational trauma and the meaning of "family continuity."

Future research should expand this model's cultural and contextual adaptability, incorporating interdisciplinary perspectives to explore how digital mourning may be personalized and ethically sensitive in AI-dominated environments. Questions worth exploring include: Do different age groups, religious backgrounds, or grief types require differentiated interfaces and commemorative modalities? Can algorithms be designed to support grief rather than standardize it? These questions touch not only on user experience optimization, but also on the moral transformation of death culture in the age of artificial intelligence. Ultimately, AI-based commemoration is not a neutral extension of human emotion, but a complex technological force that intervenes in subjectivity, ethical judgment, and cultural meaning.

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 authors.

Ethics statement

The studies involving humans were approved by the Human Research Ethics Committee for Non-Clinical Faculties of Chengdu Neusoft University (Approval No. [CNU20241120]) and was in compliance with Chinas Personal Information Protection Law and institutional data governance standards. The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation in this study was provided by the participants legal guardians/next of kin.

Author contributions

KF: Conceptualization, Investigation, Supervision, Writing – review & editing, Writing – original draft. CY: Conceptualization, Formal analysis, Investigation, Writing – original draft, Writing – review & editing. ZW: Data curation, Investigation, Software, Validation, Writing – original draft, Writing – review & editing. MW: Writing – review & editing, Conceptualization. ZL: Writing – review & editing, Formal analysis, Visualization. YY: Funding acquisition, Project administration, Validation, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This work was supported by the BNBU Center for Creative Media and Communication Research (grant number UICA0100014), and the Guangdong Province's "Breakthrough, Supplementation and Strengthening" UIC New Media Communication Teaching and Research Innovation Team (grant numbers UICR0400013-24).

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.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

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

1. Elder A. Conversation from beyond the grave? A Neo-Confucian ethics of chatbots of the dead. *J Appl Philos.* (2020) 37:73–88. doi: 10.1111/japp.12369

2. Pizzoli S, Monzani D, Vergani L, Sanchini V, Mazzocco K. From virtual to real healing: a critical overview of the therapeutic use of virtual reality to cope with mourning. *Current Psychology (New Brunswick, N.J.).* (2021) 42:8697–704. doi: 10.1007/s12144-021-02158-9

3. Krueger J, Osler L. Communing with the dead online: chatbots, grief, and continuing bonds. *J Conscious Stud.* (2022) 29(9–10):222–52. doi: 10.53765/20512201.29.9.222

4. Altaratz D, Morse T. Digital seance: fabricated encounters with the dead. *Soc Sci.* (2023) 12(11):635. doi: 10.3390/socsci12110635

5. Morris MR, Brubaker JR. Generative ghosts: anticipating benefits and risks of AI afterlives. *In Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems* (2025). p. 1–14. doi: 10.1145/3706598.3713758

6. Rodríguez Reséndiz H, Rodríguez Reséndiz J. Digital resurrection: challenging the boundary between life and death with artificial intelligence. *Philosophies.* (2024) 9(3):71. doi: 10.3390/philosophies9030071

7. Figueroa-Torres M. Affection as a service: ghostbots and the changing nature of mourning. *Computer Law & Security Review.* (2024) 52:105943. doi: 10.1016/j.clsr.2024.105943

8. Gao Y, Fan X, Hua C, Zheng H, Cui Y, Li Y, et al. Failure patterns for thymic carcinoma with completed resection and postoperative radiotherapy. *Radiother Oncol.* (2022) 178:109438. doi: 10.1016/j.radonc.2022.109438

9. Laryionava K, Pfeil T, Dietrich M, Reiter-Theil S, Hiddemann W, Winkler E. The second patient? Family members of cancer patients and their role in end-of-life decision making. *BMC Palliat Care.* (2018) 17:29. doi: 10.1186/s12904-018-0288-2

10. Russell M, Manley H. Sanctioning memory: changing identity. Using 3D laser scanning to identify two 'new' portraits of the emperor nero in English antiquarian collections. *Internet Archaeol.* (2016) 42. doi: 10.11141/IA.42.2

11. Lee P, Kim I, Yoon D. Speculating on risks of AI clones to selfhood and relationships: doppelganger-phobia, identity fragmentation, and living memories. *Proc ACM Hum Comput Interact.* (2023) 7:1–28. doi: 10.1145/3579524

12. Puzio A. When the digital continues after death: ethical perspectives on death tech and the digital afterlife. *Communicatio Socialis.* (2023) 56(3):427–36. doi: 10.5771/0010-3497-2023-3-427

13. Lu H. Exploring the predictors of public acceptance of artificial intelligence-based resurrection technologies. *Technol Soc.* (2024) 78:78. doi: 10.1016/j.techsoc.2024.102657

14. Lambert A, Nansen B, Arnold M. Algorithmic memorial videos: contextualising automated curation. *Mem Stud.* (2018) 11:156–71. doi: 10.1177/1750698016679221

15. DeBerry-Spence B, Trujillo-Torres L. Don't give US death like this! commemorating death in the age of COVID-19. *J Assoc Consum Res.* (2020) 7:27–35. doi: 10.1086/711832

16. Stahl B, Timmermans J, Mittelstadt B. The ethics of computing. *ACM Comput Surv.* (2016) 48:1–38. doi: 10.1145/2871196

17. Johnson D. What is the relationship between computer technology and ethical issues? *Interdiscip Sci Rev.* (2021) 46:430–9. doi: 10.1080/03080188.2020.1838212

18. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. *MIS Q.* (2003) 27:425–78. doi: 10.2307/30036540

19. Dwivedi YK, Rana NP, Jeyaraj A, Clement M, Williams MD. Re-examining the unified theory of acceptance and use of technology (UTAUT): towards a revised theoretical model. *Inf Syst Front.* (2019) 21:719–34. doi: 10.1007/s10796-017-9774-y

20. Alalwan AA, Baabdullah AM, Rana NP, Tamilmani K, Dwivedi YK. Examining adoption of mobile internet in Saudi Arabia: extending TAM with perceived enjoyment, innovativeness and trust. *Technol Soc.* (2018) 55:100–10. doi: 10.1016/j.techsoc.2018.06.007

21. Oliveira T, Thomas M, Baptista G, Campos F. Mobile payment: understanding the determinants of customer adoption and intention to recommend the technology. *Comput Human Behav.* (2016) 61:404–14. doi: 10.1016/j.chb.2016.03.030

22. Beldad AD, Hegner SM. Expanding the technology acceptance model with the inclusion of trust, social influence, and health valuation to determine the predictors of German users' willingness to continue using a fitness app: a structural equation modeling approach. *Int J Hum Comput.* (2018) 34(9):882–93. doi: 10.1080/10447318.2017.1403220

23. Venkatesh V, Thong JYL, Xu X. Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology. *MIS Q.* (2012) 36:157–78. doi: 10.2307/41410412

24. Kaur N, Dwivedi D, Arora J, Gandhi A. Study of the effectiveness of e-learning to conventional teaching in medical undergraduates amid COVID-19 pandemic. *Natl J Physiol Pharm Pharmacol.* (2020) 10(7):563–7. doi: 10.1145/3555803

25. Kahnbach L, Hase A, Kuhl P, Lehr D. Explaining primary school teachers' intention to use digital learning platforms for students' individualized practice: comparison of the standard UTAUT and an extended model. *Front Educ*. (2024) 9:1353020. doi: 10.3389/feduc.2024.1353020

26. Hassan I, Murad M, El-Shekeil I, Liu J. Extending the UTAUT2 model with a privacy Calculus model to enhance the adoption of a health information application in Malaysia. *Informatics*. (2022) 9:31. doi: 10.3390/informatics9020031

27. Zeebaree M, Agoyi M, Aqel M. Sustainable adoption of E-government from the UTAUT perspective. *Sustainability*. (2022) 14(9):5370. doi: 10.3390/su14095370

28. Rahi S, Othman Mansour MM, Alghizzawi M, Alnaser FM. Integration of UTAUT model in internet banking adoption context: the mediating role of performance expectancy and effort expectancy. *J Res Interact Mark*. (2019) 13(3):411–35. doi: 10.1108/jrim-02-2018-0032

29. Al-Saedi K, Al-Emran M, Ramayah T, Abusham E. Developing a general extended UTAUT model for M-payment adoption. *Technol Soc*. (2020) 62:101293. doi: 10.1016/j.techsoc.2020.101293

30. Chen G, Fan J, Azam M. Exploring artificial intelligence (AI) chatbots adoption among research scholars using unified theory of acceptance and use of technology (UTAUT). *J Librariansh Inf Sci*. (2024). doi: 10.1177/09610006241269189

31. Li M, Fu D, Lin L, Cao Y, He H, Siyu Z, et al. Pathway analysis of clinical nurse educator's intention to use virtual reality technology based on the UTAUT model. *Front Public Health*. (2024) 12:1437699. doi: 10.3389/fpubh.2024.1437699

32. De Luna I, Jiménez-Alonso B. Deathbots. Discussing the use of artificial intelligence in grief/ deathbots. Debatiendo el uso de la inteligencia artificial en el duelo. *Stud Psychol*. (2024) 45:103–22. doi: 10.1177/02109395241241387

33. Allison A, Currall J, Moss M, Stuart S. Digital identity matters. *J Am Soc Inf Sci Technol*. (2005) 56(4):364–72. doi: 10.1002/asi.20112

34. Nakagawa H, Orita A. Using deceased people's personal data. *AI Soc*. (2024) 39(3):1151–69. doi: 10.1007/s00146-022-01549-1

35. Ienca M. On artificial intelligence and manipulation. *Topoi (Dordr)*. (2023) 42(3):833–42. doi: 10.1007/s11245-023-09940-3

36. . Foucault M., Mailander E. Surveiller et punir (Vol. 225). na. (1975).

37. Eriksson Krutrök M. Algorithmic closeness in mourning: vernaculars of the hashtag #grief on TikTok. *Social Med Soc*. (2021) 7(3). doi: 10.1177/20563051211042396

38. Capodivacca S, Giacomini G. Discipline and power in the digital age: critical reflections from foucault's thought. *Foucault Stud*. (2024) 36(1):227–51. doi: 10.22439/fs.i36.7215

39. Brubaker J, Hayes G, Dourish P. Beyond the grave: facebook as a site for the expansion of death and mourning. *Inf Soc*. (2013) 29:1152–163. doi: 10.1080/01972243.2013.777300

40. Myles D, Cherba M, Millerand F. Situating ethics in online mourning research: a scoping review of empirical studies. *Qual Inq*. (2019) 25:289–99. doi: 10.1177/1077800418806599

41. Harbinja E, Edwards L, McVey M. Governing ghostbots. *Comput Law Secur Rev*. (2023) 48:105791. doi: 10.1016/j.clsr.2023.105791

42. Wang X, Li X, Yin Z, Wu Y, Liu J. Emotional intelligence of large language models. *J Pac Rim Psychol*. (2023) 17:18344909231213958. doi: 10.1177/18344909231213958

43. Widmaier L. Digital photographic legacies, mourning, and remembrance: looking through the eyes of the deceased. *Photographies*. (2023) 16:19–48. doi: 10.1080/17540763.2022.2150879

44. Lietor P, Penichet V, Lozano M. Emotional interaction in social networks. *Eur Psychiatry*. (2022) 65(S1):S577–S577. doi: 10.112/j.eurpsy.2022.1478

45. El-Jawahri A, Greer J, Park E, Jackson V, Kamdar M, Rinaldi S, et al. Psychological distress in bereaved caregivers of patients with advanced cancer. *J Pain Symptom Manage*. (2020) 61(3):488–94. doi: 10.1016/j.jpainsymman.2020.08.028

46. El-Jawahri A, Traeger L, Park ER, Greer JA, Pirl WF, Jackson VA, et al. Depression and anxiety symptoms in bereaved caregivers of patients with advanced cancer. *J Clin Oncol*. (2019) 37(Suppl 31):17. doi: 10.1200/JCO.2019.37.31_suppl.17

47. Cassileth BR, Lusk EJ, Strouse TB, Miller DS, Brown LL, Cross PA. A psychological analysis of cancer patients and their next-of-kin. *Cancer*. (1985) 55(1):72–6. doi: 10.1002/1097-0142(19850101)55:1<72::AID-CNCR2820550112>3.0.CO;2-S

48. Pitceathly C, Maguire P. The psychological impact of cancer on patients' partners and other key relatives: a review. *Eur J Cancer*. (2003) 39(11):1517–24. doi: 10.1016/S0959-8049(03)00309-5

49. Hermann M, Goerling U, Hearing C, Mehnert-Theuerkauf A, Hornemann B, Hvel P, et al. Social support, depression and anxiety in cancer patient-relative dyads in early survivorship: an actor-partner interdependence modeling approach. *Psychooncology*. (2024) 33:12. doi: 10.1002/pon.70038

50. Rosenberg A, Postier A, Osenga K, Kreicbergs U, Neville B, Dussel V, et al. Long-term psychosocial outcomes among bereaved siblings of children with cancer. *J Pain Symptom Manage*. (2015) 49(1):55–65. doi: 10.1016/j.jpainsymman

51. Rosner R, Pfoh G, Kotoucová M. Treatment of complicated grief. *Eur J Psychotraumatol*. (2011) 2(1):7995. doi: 10.3402/ejpt.v2i0.7995

52. Golden A, Dalgleish T. Is prolonged grief distinct from bereavement-related posttraumatic stress? *Psychiatry Res*. (2010) 178:336–41. doi: 10.1016/j.psychres.2009.08.021

53. Melhem N, Day N, Shear M, Day R, Reynolds C, Brent D. Traumatic grief among adolescents exposed to a peer's suicide. *Am J Psychiatry*. (2004) 161(8):1411–6. doi: 10.1176/appi.ajp.161.8.1411

54. Rosner R, Comtesse H, Vogel A, Doering B. Prevalence of prolonged grief disorder. *J Affect Disord*. (2021) 287:301–7. doi: 10.1016/j.jad.2021.03.058

55. De Stefano R, Muscatello M, Bruno A, Cedro C, Mento C, Zoccali R, et al. Complicated grief: a systematic review of the last 20 years. *Int J Soc Psychiatry*. (2020) 67:492–9. doi: 10.1177/0020764020960202

56. Prigerson HG, Kakarala S, Gang J, Maciejewski PK. History and status of prolonged grief disorder as a psychiatric diagnosis. *Annu Rev Clin Psychol*. (2021) 17(1):109–26. doi: 10.1146/annurev-clinpsy-081219-093600

57. Lichtenthal WG, Catarozoli C, Masterson M, Slivjak E, Schofield E, Roberts KE, et al. An open trial of meaning-centered grief therapy: rationale and preliminary evaluation. *Palliat Support Care*. (2019) 17(1):2–12. doi: 10.1017/S1478951518000925

58. Raghunathan R, Pham M. All negative moods are not equal: motivational influences of anxiety and sadness on decision making. *Organ Behav Hum Decis Process*. (1999) 79(1):56–77. doi: 10.1006/obhd.1999.2838

59. Suo T, Jia X, Song X, Liu L. The differential effects of anger and sadness on intertemporal choice: an ERP study. *Front Neurosci*. (2021) 15:638989. doi: 10.3389/fnins.2021.638989

60. Garg N, Lerner J. Sadness and consumption. *J Consum Psychol*. (2013) 23:106–13. doi: 10.1016/j.jcps.2012.05.009

61. Lei P, Zhang H, Zheng W, Zhang L. Does sadness bring myopia: an intertemporal choice experiment with college students. *Front Psychol*. (2024) 15:1345951. doi: 10.3389/fpsyg.2024.1345951

62. Eisma MC. Prolonged grief disorder in ICD-11 and DSM-5-TR: challenges and controversies. *Aust N Z J Psychiatry*. (2023) 57(7):944–51. doi: 10.1177/00048674231154206

63. She WJ, Ang CS, Neimeyer RA, Burke LA, Zhang Y, Jatowt A, et al. Investigation of a web-based explainable AI screening for prolonged grief disorder. *IEEE Access*. (2022) 10:41164–85. doi: 10.1109/ACCESS.2022.3163311

64. Ismatullaev U, Kim S. Review of the factors affecting acceptance of AI-infused systems. *Hum Factors*. (2022) 66:126–44. doi: 10.1177/00187208211064707

65. Botella C, Osma J, Palacios A, Guillén V, Baios R. Treatment of complicated grief using virtual reality: case report. *Death Stud*. (2008) 32:674–92. doi: 10.1080/07481180802231319

66. Sas C, Ren S, Coman A, Clinch S, Davies N. Life review in end of life care: a practitioner's perspective. In: *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '16)*. New York, NY: Association for Computing Machinery (2016). p. 2947–53. doi: 10.1145/2851581.2892491

67. Prigerson HG, Maciejewski PK, Reynolds CF III, Bierhals AJ, Newsom JT, Fasiczka A, et al. Inventory of complicated grief: a scale to measure maladaptive symptoms of loss. *Psychiatry Res*. (1995) 59(1–2):65–79. doi: 10.1016/0165-1781(95)02757-2

68. Hair JF, Sarstedt M, Pieper TM, Ringle CM. The use of partial least squares structural equation modeling in strategic management research: a review of past practices and recommendations for future applications. *Long Range Plann*. (2012) 45(5–6):320–40. doi: 10.1016/j.lrp.2012.09.008

69. Sarstedt M, Ringle CM, Hair JF. *Partial Least Squares Structural Equation Modeling. in Handbook of Market Research*. Cham: Springer International Publishing (2021). p. 587–632. doi: 10.1007/978-3-319-57413-4_15

70. Streukens S, Leroi-Werelds S. Bootstrapping and PLS-SEM: a step-by-step guide to get more out of your bootstrap results. *Eur Manag J*. (2016) 34(6):618–32. doi: 10.1016/j.emj.2016.06.003

71. Edeh E, Lo WJ, Khojasteh J. Review of Partial Least Squares Structural Equation Modeling (PLS-SEM) Using R: A Workbook by Joseph F. Hair Jr., G. Tomas M. Hult, Christian M. Ringle, Marko Sarstedt, Nicholas P. Danks, Soumya Ray. Cham, Switzerland: Springer; 2021. *Struct Equ Modeling*. (2023) 30(1):165–7. doi: 10.1080/10705511.2022.2108813

72. Hair JF, Ringle CM, Sarstedt M. PLS-SEM: indeed a silver bullet. *J Mark Theory Pract*. (2011) 19(2):139–52. doi: 10.2753/MTP1069-6679190202

73. Henseler J, Sarstedt M. Goodness-of-fit indices for partial least squares path modeling. *Comput Stat*. (2013) 28:565–80. doi: 10.1007/s00180-012-0317-1

74. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Q*. (1989) 13(3):319–40. doi: 10.2307/249008

75. Garcia de Blanes Sebastian M, Sarmiento Guede JR, Antonovica A. Application and extension of the UTAUT2 model for determining behavioral intention factors in use of the artificial intelligence virtual assistants. *Front Psychol*. (2022) 13:993935. doi: 10.3389/fpsyg.2022.993935

76. Venkatesh V, Morris MG. Why don't men ever stop to ask for directions? Gender, social influence, and their role in technology acceptance and usage behavior. *MIS Q*. (2000) 24(1):115–39. doi: 10.2307/3250981

77. Harbinja E, Morse T, Edwards L. Digital remains and post-mortem privacy in the UK: what do users want? *Int Rev Law Comput Technol*. (2025):1–24. doi: 10.1080/13600869.2025.2506164

78. Attuquayefio S, Addo H. Using the UTAUT model to analyze students' ICT adoption. *IJEDICT*. (2014) 10(3). Available online at: https://www.learntechlib.org/p/148478/

79. Zhu W, Huang L, Zhou X, Li X, Shi G, Ying J, et al. Could AI ethical anxiety, perceived ethical risks and ethical awareness about AI influence university Students' use of generative AI products? An ethical perspective. *Int J Hum Comput*. (2024) 41(1):742–64. doi: 10.1080/10447318.2024.2323277

80. Li B, Qi P, Liu B, Di S, Liu J, Pei J, et al. Trustworthy AI: from principles to practices. *ACM Comput Surv*. (2023) 55(9):1–46. doi: 10.1145/3555803

81. Xiyi C. The construction and application of relationship law theory: a case study of re-understanding the marital fidelity agreement. *Modern Law*. (2024) 46(04):192–208. doi: 10.3969/j.issn.1001-2397.2024.04.13

82. Chao YH, Wang SY, Sheu SJ. Integrative review of breast cancer survivors' transition experience and transitional care: dialog with transition theory perspectives. *Breast Cancer*. (2020) 27(5):810–8. doi: 10.1007/s12282-020-01097-w

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 08 August 2025
DOI 10.3389/fpsyg.2025.1652130

Check for updates

OPEN ACCESS

EDITED BY
Wenjing Chen,
Beijing University of Posts and Telecommunications (BUPT), China

REVIEWED BY
Taejong Kim,
National Institute of Meteorological Sciences, Republic of Korea
Yew Chee Chew,
Malaysia University Science and Technology, Malaysia

*CORRESPONDENCE
Dongling Ma
✉ madongling@wtbu.edu.cn

RECEIVED 23 June 2025
ACCEPTED 16 July 2025
PUBLISHED 08 August 2025

CITATION
Jiang W, Li Y, Hu X and Ma D (2025) Reframing individual roles in collaboration: digital identity construction and adaptive mechanisms for resistance-based professional skills in AI-human intelligence symbiosis. *Front. Psychol.* 16:1652130. doi: 10.3389/fpsyg.2025.1652130

Reframing individual roles in collaboration: digital identity construction and adaptive mechanisms for resistance-based professional skills in AI-human intelligence symbiosis

Weizheng Jiang[1,2], Yongzhou Li[1], Xiaoling Hu[2] and Dongling Ma[2]*

[1]School of Management, Wuhan University of Science and Technology, Wuhan, China, [2]School of Management, Wuhan Technology and Business University, Wuhan, China

Amid the unprecedented wave of AI advancement, AI-resistant professional skills play a significant role in enhancing the effectiveness of human–AI collaboration. However, existing research tends to isolate professional skills from their broader context, overlooking the triadic construction of digital identity recognition through individual motivation, structural position, and knowledge articulation. This oversight weakens the sustainability and adaptability of skill expression, thereby hindering innovation performance in AI–HI (Artificial Intelligence–Human Intelligence) collaboration. Drawing on the entropy weight method, gradient descent algorithm, and a residual–matching decision matrix, this study conducted quantitative modeling of 418 participants in the financial co-production sector from 2022 to 2024. The findings reveal that network centrality (NC; $\beta = 0.04^{**}$) and proactive personality (PP; $\beta = 0.05^{**}$) significantly amplify the impact of two key AI-resistant skills—foreign language proficiency (FL) and passion/optimism (PO)—on collaboration effectiveness, through structural empowerment and intrinsic motivation. Furthermore, this study develops a digital identity recognition and classification framework that identifies three distinct groups: core innovators, marginal experts, and low performers. By extending the theoretical model of digital identity construction within AI–HI collaboration, this study also proposes a differentiated approach to talent development and resource allocation based on innovation effectiveness and identity alignment, offering new insights into the advancement of digital human capital.

KEYWORDS

digital identity construction, AI-resistant professional skills, AI-HI collaborative innovation performance, knowledge articulation, collaborative intelligence

1 Introduction

Since 2022, the development of generative AI (GenAI) has accelerated at an extraordinary pace, continuously deepening its integration and entanglement with a wide range of industries. However, evidences have emerged pointing to a set of global challenges accompanying AI's advancement: the remarkable acceleration in technological development has not been matched by a corresponding increase in collaborative effectiveness. According to the *AI Index Report 2024* (Clark et al., 2024), investments related to GenAI surged 8fold in 2023, reaching USD 25.2 billion. In specific tasks such as image classification and language comprehension, AI has already surpassed human performance. In a striking contrast, the report also reveals that in complex tasks—such as visual reasoning, mathematical reasoning, and creative problem-solving—AI has not demonstrated a clear advantage.

Furthermore, McKinsey's *The State of AI* report (Singla et al., 2025) points to a series of organizational challenges faced by companies embedding AI into their business processes: misalignment between deployment workflows and human resource structures, difficulties in hiring for AI-related roles, underdeveloped governance and trust mechanisms, and the need for a fundamental reconfiguration of AI-HI interaction. Collectively, these conditions underscore the necessity of AI-human symbiosis, while also revealing that effective mechanisms for AI-HI collaboration have yet to be fully established.

Echoing this global concern, China is likewise experiencing a "high technology-low collaboration" paradox. According to the CAICT (2024), China's core AI industry grew by 13.9% year-on-year in 2023, and by the first quarter of 2024, a total of 478 large-scale AI models had been released. However, data from the iResearch Group (2024) indicates that the growth of China's collaborative office platforms has slowed significantly since 2022 and has even begun to decline. This structural tension between rapid technological advancement and declining collaboration raises fundamental questions about AI-HI interaction: does the acceleration of AI development necessarily lead to collaborative innovation at the organizational level? Or does the disjunction between speed and effectiveness reflect a deeper disruption—one in which AI challenges not only the traditional roles of human collaboration but also the cognitive mechanisms underlying them? In such an AI-human symbiotic environment, must humans seek new forms of identity and positioning?

This dilemma is particularly evident in practical applications. For example, Zoom's *AI Companion*, regarded as a benchmark for remote collaboration, achieves only about a 50% success rate in generating accurate meeting summaries. Although generative AI (GenAI) possesses high programmability and strong logical capabilities, the absence of mechanisms for "role construction" and "digital identity adaptation" often results in information silos, unclear responsibility boundaries, and impeded collaboration (Okamura and Yamada, 2020). Moreover, the challenge of task allocation between humans and AI is deeply shaped by the characteristics of the AI system itself. A lack of transparency or limited perceptibility can directly undermine trust mechanisms in human–AI interactions (Rai, 2020). In addition, low levels of anthropomorphism in AI often result in an absence of emotional support (Sheng et al., 2024). Overlooking these system-level factors can lead to role misalignment and psychological dissonance within collaborative human–AI settings.

These findings prompt a rethinking of the practical challenges in AI-HI collaborative innovation. By approaching the issue through the lens of individual digital identity construction, this study identifies mismatch patterns in identity recognition that affect collaboration performance and proposes a mechanism to optimize identity-role alignment.

Existing research widely emphasizes the value of AI in collaborative innovation, highlighting its advantages in prediction, efficiency, and knowledge acquisition (Chu et al., 2023). However, much of the literature remains trapped in a binary opposition between technological determinism and capability determinism: the former overemphasizes the direct impact of technological features, while the latter treats professional skills in isolation from their broader context. Both perspectives neglect the complexity inherent in establishing a symbiotic relationship between AI-HI. As illustrated in McKinsey's *What Employees Are Saying About the Future of Remote Work* report (Andrea et al., 2021), factors such as social network construction and psychological needs play a critical role in digital workplace collaboration. In fact, there is a complex and dynamic interplay among individual intrinsic motivations, social network positions, knowledge articulation, and AI-HI role allocation (Alowais et al., 2023; Jia et al., 2024). Only by considering and modeling these elements holistically can we effectively uncover the true mechanisms underlying AI-HI collaborative innovation.

Building on the above background, this study seeks to address a critical question: how do professional skills, coexisting with AI, influence collaborative innovation in multifaceted environments? To this end, we adopt a triadic synergy perspective that emphasizes the interaction among individual personality traits, the distribution of professional skills, and social network construction, and develop a quantitative framework for analyzing AI-HI collaborative innovation.

However, existing research on AI-driven professional skills exhibits three main limitations. First, many studies focus solely on the distribution of professional skills within industries, overlooking the actual impact of these skills on AI-enabled innovation performance (e.g., Alekseeva et al., 2021; Fletcher and Thornton, 2023). Second, few investigations integrate personal traits and social networks, resulting in overfitting issues between skills and behavioral effects. For instance, Chuang (2024) argues that the full potential of AI technologies depends on collective intelligence. Third, social network position disrupts the pathways of professional skill conversion. This is especially true for soft skills, whose effectiveness often hinges on informal networks—an aspect insufficiently addressed in current literature (Cangialosi et al., 2021; e.g., Burtch et al., 2023).

From a practical perspective, existing research remains insufficient in decomposing and prioritizing AI-related professional skills. Quantitative assessments of such skills are largely based on self-reported questionnaires, which lack objectivity. Hard skills, being observable and recordable capabilities, should be measured through more rigorous and objective approaches (i.e., Alekseeva et al., 2021). Moreover, the extent to which professional skills are translated into collaborative innovation performance should not be inferred solely from subjective scales (e.g., Cangialosi et al., 2021; Chowdhury et al., 2022; Pham et al., 2024); instead, they should be evaluated in relation to the actual task allocation between humans and AI within the collaboration process.

To address these limitations and research gaps, we implemented the following measures. First, we incorporated a triadic synergy framework to investigate the interaction among personality traits, social network structure, and professional skill expression in AI-HI collaborative innovation. Second, we adopted the After Action Review (AAR) method to objectively capture cognitive performance outcomes and assess the effectiveness of AI-HI collaboration, particularly with regard to hard skills. Third, we applied Gradient Descent (GD) techniques to explore the distribution of AI-driven professional skills, rather than treating soft and hard skills as monolithic constructs, thereby improving the sensitivity of measurement and prioritization. Fourth, we constructed a residual- matching analytical matrix to reconceptualize digital identity from an interactional perspective and proposed targeted knowledge management strategies accordingly.

This study surveyed 418 participants from AI-enabled co-production industries in the financial sector between 2022 and 2024. Data were collected on proactive personality (PP) traits, community network structures, and cognitive skills to uncover the mechanisms through which AI-driven professional skills shape AI-HI role allocation. These findings offer new insights into the formation and development of digital human capital theory.

2 Theoretical framework and research questions

2.1 Digital identity construction in triadic synergy

2.1.1 The definition of digital identity and essential dimensions

The construction of digital identity spans multiple disciplines, including information technology, sociology, and psychology. Its formation is dynamic, continuous, and multilayered. According to Sedlmeir et al. (2021), digital identity refers to the representation of an entity within a virtual environment—an entity that may include not only individual humans but also legal persons or technological agents. A digital identity is composed of controllable attributes (such as behavioral records, technical features, and social parameters) and is anchored by a unique identifier that ensures consistency across platforms. This definition emphasizes both the sustainability (recognizability) and the diversity of identity features. Furthermore, digital identity is co-constructed by individuals, organizations, and technological systems. With the emergence of decentralized management approaches, identity assignment is increasingly distributed and individually governed, granting users greater autonomy and digital sovereignty.

Based on the definition and connotation of digital identity, this study proposes a triadic synergy mechanism composed of intrinsic motivation, structural positioning, and knowledge articulation. This framework aims to capture the generative logic of digital identity in an AI-HI symbiotic environment. The construction of such identity is not only critical to the allocation of roles in AI-HI collaboration but also plays a key role in the sustainable adaptation of knowledge co-creation and innovation.

Digital identity within the triadic synergy framework comprises three essential dimensions:

2.1.1.1 Knowledge articulation

Represented by quantifiable professional skills, this dimension reflects the extent to which individuals can mobilize their intrinsic motivations and leverage structural positions in AI-HI collaboration (e.g., Chuang, 2024). Collaborative intelligence emphasizes the complementarity, reciprocity, and co-evolution between humans and AI. It reflects not only the technical challenges of embedding AI into work processes, but also the alignment and synergy between AI systems and individual knowledge articulation (Tariq et al., 2025).

2.1.1.2 Structural positioning

Measured through social network centrality, this captures the external visibility and structural opportunities of one's identity on digital platforms (i.e., Morrison-Smith and Ruiz, 2020). This indicator directly pertains to hybrid teaming frameworks—specifically, how structural position advantages can be leveraged to integrate human creativity and emotional intelligence with the strengths of AI, thereby forming a highly efficient and adaptable collaborative unit (Caldwell et al., 2022).

2.1.1.3 Intrinsic motivation

Exemplified by PP traits, this dimension provides the motivational foundation of digital identity. In AI-HI collaboration, self-determination positively influences exploratory activities in AI-HI symbiosis (e.g., Kong et al., 2024). When AI is designed as human-centered (or HCAI) and collaboration-enhancing, individuals are more likely to exhibit proactive agency rather than passive acquiescence or substitution anxiety (Shneiderman, 2022).

2.1.2 Knowledge articulation: professional skills

2.1.2.1 Hard skills in AI-HI collaboration

According to Hendarman and Cantner (2018), hard skills refer to professional knowledge and technical competencies that can be described, quantified, preserved, and documented, and these sets of competencies are relevant to specific tasks. The content of hard skills varies across professional fields, but the vast majority of hard skill sets involve computer competence and digital literacy, as investigated by Alekseeva et al. (2021), and Chuang (2024).

In the integration of computer and AI technologies, computer operations, machine learning understanding, and programming involve the expression of computer and network technologies in individual behavior, i.e. Cyber Behavior (CB). Analysis, modeling, and mathematical foundations affect the individual's access to the laws behind the data, i.e., Date Analysis (DA). They play a key role in the understanding and application of AI. Not only do they reflect learners' mastery of superficial AI skills, but also demonstrate their understanding and insight into the principles and mechanisms behind the operation of AI, as Bankins et al. (2024) conclude that algorithms and human capabilities influence employees' own experience and job design.

However, the differential impact from hard skills is not only reflected in job and innovation performance, but also in the regional adaptation to digital technologies, as investigated by Carlisle et al. (2023), who argues that CB and DA are far less important than communication skills in the digital transformation of the service sector. Whereas, mastery of a Foreign Language

(FL) has been considered as an important hard skill in many studies, functioning to facilitate cross-cultural communication and knowledge flow (Zeng and Yang, 2024).

2.1.2.2 Soft skills in AI-HI collaboration

According to Hendarman and Cantner (2018) summary, soft skills involve relational resources and communication skills aimed at environmental adaptation through interpersonal embedding and are informal skills. In contrast to hard skills, the prominent role of soft skills lies in the development of communication skills, teamwork and problem solving skills. For human-computer coexistence relationships, these competences are not only necessary for the modern workplace (Escolà-Gascón and Gallifa, 2022). Their acquisition and enhancement can also increase human-AI trust, playing a key role in the maintenance of friendly AI-HI coexistence relationships (Sheng et al., 2024).

Soft skills can be expressed as Innovation Leadership (IL), Relationship Building (RB), Tolerance for Uncertainty (TU), and Passion and Optimism (PO) (2018). Soft skills likewise drive digital transformation and demonstrate criticality in AI-driven collaborative innovation. IL is related to technology acceptance management, which influences collaboration across disciplines and domains (Bahoo et al., 2023). RB competencies derive from field theory and group dynamics, which are beneficial for the building of hybrid AI social networks (Ng, 2022). Leadership types holding TUs are beneficial for promoting psychological safety and facilitating incremental innovation (Uhl-Bien, 2021), and PO sustains positive individual perceptions of change and coexistence in the AI era (Burtch et al., 2023).

Therefore, as a crucial component of the knowledge articulation, professional skills carry the dual function of task execution and knowledge exchange within AI–HI collaboration. Hard skills reflect an individual's depth of understanding and operational proficiency with AI technologies, while soft skills signify the capacity to collaborate effectively, enhance communication, and foster trust. Together, these skill sets shape the visibility, reliability, and role recognition of individuals in the process of digital identity construction.

2.1.3 Structural position: network centrality

Social capital is manifested through structural positions within social networks, with centrality (including relative centrality, betweenness centrality, etc.) being a key indicator of social capital. Individuals occupying central positions, with higher information flow efficiency and influence, can more effectively promote knowledge spillover and innovation (Cangialosi et al., 2021).

Building on this foundation, hybrid AI social networks not only optimize information transmission pathways and strengthen relational connections. These enhancing individuals' structural embeddedness, but also amplify the leverage effect of personal centrality on knowledge collaboration through augmented social computing. As Wang et al. (2022) argue, H-AI (hybrid human–AI) systems, as an augmented intelligence paradigm, are essential in addressing the limitations of conventional AI when confronting complex and dynamic social problems. This perspective indirectly highlights the foundational role that network structural advantages play in the construction of social behavior within digital identity formation.

From the perspective of hard skills, digital capability gaps often lead to issues in task completion time and efficiency. These gaps are heavily influenced by one's structural position within the social network. Individuals at the core of the network typically have greater access to information and communication opportunities, which in turn accelerates the expression and application of hard skills (Lythreatis et al., 2022).

Soft skills can be expressed as communication abilities, teamwork capabilities, adaptability, and emotional intelligence. These skills are integral to the maintenance and establishment of centrality. Furthermore, while virtual teams and remote collaboration rely on technology, communication and trust remain essential. Particularly in virtual networks, self-organization can easily lead to disorderly project development. Therefore, a higher Network Centrality (NC) implies trust and commitment, which facilitates the correct and efficient operation of professional skills.

In summary, NC reflects two key aspects of an individual's structural position: the efficiency of information flow and knowledge collaboration, and the recognition of one's role within their digital identity. Within hybrid AI-enabled social networks, individuals with higher centrality are more likely to be identified as knowledge hubs. Such structural advantages enhance the external visibility and system-level recognition of their digital identity, making their role in AI–HI collaboration more stable and sustainable.

2.1.4 Intrinsic motivation: proactive personality

Proactive Personality (PP), as a form of intrinsic motivation, is closely linked to self-determination. It is often studied in terms of how individuals activate themselves positively by constructing both internal and external sources of meaning. First, from the perspective of external relationship building, proactive individuals initiate positive changes in their interactions through internal drives, expanding their social networks and thereby increasing the likelihood of realizing personal visions (Rienda et al., 2025). Second, from the standpoint of internal meaning-making, individuals with high levels of proactivity tend to assign deeper significance to learning, work, and personal effort. This capacity enables them to more effectively translate intrinsic motivation into concrete expressions of creativity (Zhang et al., 2021).

From the perspective of technology acceptance theory, trust in AI is associated with dimensions such as transparency, functionality, reliability, explainability, and perceived benevolence (Huang and Rust, 2024). While transparency, explainability, and functionality relate to the design and performance of AI systems, the extent to which these features are perceived and trusted largely depends on users' willingness to engage and their cognitive orientation. Zheng et al. (2020), through a study on online learning during the COVID-19 pandemic, found that PP significantly enhanced the quality of online interaction, self-efficacy, and social capital. Our latest research also indicates that excessive AI transparency may hinder the development of higher-order cognition. Enhancing individual digital identity through collaborative intelligence depends on the cultivation of effective learning and self-management abilities (Jiang et al., 2025). These findings suggest that proactive individuals are more inclined to explore and engage with technology, which in turn improves their understanding of and trust in the complex mechanisms underlying AI—thus forming a psychological foundation for technology trust.

Therefore, as a driving force behind digital identity formation, a proactive "self-driven identity" not only facilitates skill expression and network embeddedness in AI–HI collaboration but also strengthens the trust relationship between humans and AI systems.

2.2 Limitations in measuring AI-HI collaborative innovation and cognitive dynamics

The dynamic reorganization of cognitive resources within social networks is the essence of knowledge collaboration and innovation. Nonaka's (2022) early SECI model indirectly reflects the phased progression of cognitive strategies: observational imitation (socialization) → metaphorical encoding (externalization) → deconstruction and reconstruction (combination) → learning transfer (internalization). Bloom's taxonomy further explains the transformation of individual cognition in the learning process. By defining cognitive levels, Bloom's framework provides an operational theoretical label for quantifying knowledge collaboration and innovation. Logically, this framework aligns with the SECI knowledge spiral model (see Figure 1). The socialization stage, which relies on observational imitation, corresponds to Bloom's cognitive categories of remembering and understanding, while the combination stage, requiring deconstruction and reconstruction, aligns with application and analysis. These cognitive behaviors form the fundamental units of the innovation process, and the distribution of different cognitive levels within this process serves as an indicator of collaboration and innovation efficiency.

However, the staged division of cognitive levels remains ambiguous, and AI intervention is reshaping the division of cognitive labor in the following ways:

2.2.1 AI as a substitute for lower-order cognition

AI automates a vast range of procedural and repetitive cognitive tasks, such as encoding, decoding, and information retrieval (corresponding to the externalization and combination stages), thereby creating favorable conditions for the acquisition of complex and breakthrough cognition (e.g., Hu et al., 2021).

2.2.2 AI as an enhancer of higher-order cognition

While offloading large volumes of procedural and repetitive tasks, collaborative intelligence facilitates the transfer and conversion of complex tacit knowledge—thus accelerating internalization (Asrifan et al., 2024). At the same time, human-centered AI and generative technologies can effectively support decision-making and foster creativity (Shneiderman, 2022; Jia et al., 2024; Ritala et al., 2024).

Interestingly, this disruptive transformation reveals a dual deficiency in current quantitative research approaches:

2.2.3 Measurement distortion

Current studies rely heavily on self-reported surveys (e.g., Pham et al., 2024), which fail to capture AI-induced variations in cognitive behavior, such as the flow and conversion of tacit knowledge in AI-HI collaboration. Consequently, they struggle to effectively distinguish AI-enabled collaborative innovation from traditional modes.

2.2.4 Paradigmatic lag

Existing cognitive quantification research continues to conceptualize cognitive levels as a linear progression, neglecting the "cognitive leaps" and transformative shifts induced by AI. For instance, Bharatha et al. (2024) argue that ChatGPT functions as a complementary mechanism to human cognition in medical education, reinforcing the notion that AI fundamentally reshapes human cognitive processes.

Therefore, it is necessary to establish a new paradigm for cognitive quantification to reflect the efficiency driven by AI. Mainly through the After-Action Review (AAR; Keiser and

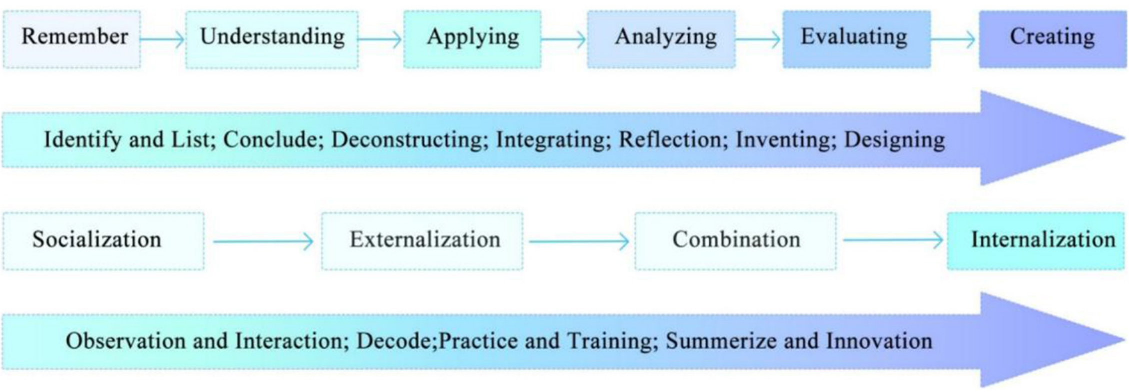


FIGURE 1
Correlation between SECI model and Bloom's Taxonomy.

Arthur, 2021), individual cognition is quantified, and the cognitive changes and reorganization during knowledge transformation are dynamically tracked (Nonaka and Yamaguchi, 2022; Asrifan et al., 2024) to reveal the laws of AI-HI collaboration.

2.3 Research questions

Building upon the triadic synergy framework, this study conceptualizes digital identity as the outcome of interactions among individual role cognition (i.e., PP), platform-based recognition (i.e., social network position), and knowledge articulation (i.e., professional skills) in AI-HI collaboration. An individual's identity is formed through the dynamic interplay of these three components, which together shape the cognitive impact of AI-HI collaboration. This dynamic mechanism and its effect on collaborative innovation effectiveness are illustrated in Figure 2.

Based on this framework, we pose the following research questions:

> RQ1: Which professional skills demonstrate greater "AI resistance," meaning they are more likely to consistently convey identity and knowledge value in collaborative settings?
> RQ2: Within the AI-HI symbiotic relationship, can the triadic synergy mechanism effectively reflect the process of digital identity construction?
> RQ3: Does an individual's position within a social network serve to empower or marginalize their digital identity? Does it amplify or suppress knowledge articulation?
> RQ4: Is there a misalignment between individuals' self-perceived identity and the system's role recognition? If so, how does this misalignment affect AI-HI collaboration?

3 Research methods

3.1 Research design

3.1.1 Measurement methods for knowledge conversion

The entropy weight method aims to construct a weight model that describes and focuses on the distribution and uncertainty of system information. The size of the entropy value is inversely proportional to the amount of information: the higher the information content of an indicator, the greater its weight. This method is primarily used for objectively analyzing internal system changes and the relationships between indicators (Zhu et al., 2020).

The entropy weight method is widely applied in the field of management to provide auxiliary weight coefficients for decision-making, thereby improving the scientific and objective nature of decisions (Chen, 2021). Additionally, this method can be applied to management and optimization across various domains (Zhu et al., 2020). Its advantages lie in reducing human interference, enhancing applicability, efficiency, and predictability.

Weight data are typically obtained using methods such as Analytic Hierarchy Process (AHP) and Principal Component Analysis (PCA). However, AHP is subject to strong subjectivity, inconsistent ratings, and varying applicability, while PCA is primarily used for variable selection and may not accurately reflect the relative weights of variables. This study innovatively uses the entropy weight method to construct a cognitive level quantification model for collaborative innovation, with the following advantages:

3.1.1.1 Objective weight assignment

Based on the principle of information entropy, this method avoids human interference and precisely quantifies the contribution weights of different cognitive levels in knowledge transformation.

3.1.1.2 System adaptation

It is suitable for the cognitive interaction relationships in AI-driven innovation and supports the analysis of the non-linear collaborative effects of variables.

However, the entropy weight method is sensitive to data quality and is limited by linear assumptions (Gray, 2011). Therefore, this study implements a three-tier optimization process:

3.1.1.3 Data preprocessing

Abnormal values are removed through validity tests and correlation analysis.

3.1.1.4 Normalization adjustment

Data are standardized during the collection process and non-dimensionalized during the analysis phase.

3.1.1.5 Model validation

By integrating the theoretical review from Section 2.2, we confirm whether the weight distribution aligns with the disruptive impact of AI on cognition.

First, the relevant indicators are standardized (see Equation 1):

$$x_{ij} = \frac{x_{ij} - \min(x_j)}{\max(x_j) - \min(x_j)} \tag{1}$$

Here, x_{ij} represents the i-th indicator of sample j .

Next, the entropy value for each indicator is established (see Equation 2), among others $p_{ij} = \frac{z_{ij}}{\sum_{j=1}^{n} z_{ij}}$, where n represents the sample size.

$$e_j = -\frac{1}{\ln(n)} \sum_{i=1}^{n} p_{ij} \ln(p_{ij}) \tag{2}$$

Finally, the weight for each indicator is calculated (see Equation 3), where m represents the total number of indicators.

$$w_j = \frac{1 - e_j}{\sum_{j=1}^{m} (1 - e_j)} \tag{3}$$

Based on Equations 1–3, a knowledge conversion effectiveness model (Equation 4) established on the weights of cognitive levels can be derived. In this framework, the AI-HI knowledge conversion outcome (KC) represents the effectiveness of AI-HI collaboration. K_j represents the acquisition effectiveness of different types of cognitive levels, b represents the intercept, and ω_j represents the weights of different cognitions.

$$KC = w_j \cdot \sum K_j + b \tag{4}$$

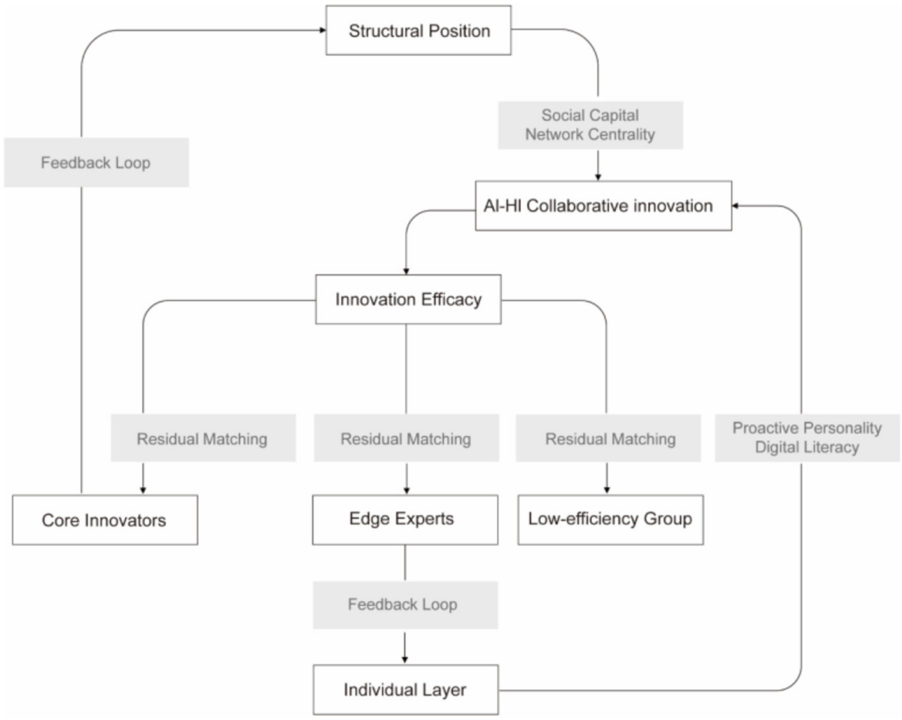


FIGURE 2
Research model and conceptual framework.

3.1.2 Constructing the digital identity recognition function from a triadic synergy perspective

This study employs mathematical modeling to represent the interactive influence among the three dimensions of triadic synergy, capturing the pathway through which individual digital identities are constructed in AI-HI collaboration. As illustrated in Figure 2, an individual's digital identity can be reflected through the effectiveness of AI-HI knowledge conversion.

To capture the sensitivity of triadic synergy variables—particularly the distribution and ranking of professional skills—we adopt the Gradient Descent (GD) algorithm to model multidimensional non-linear relationships. As a core optimization technique in machine learning, GD derives the minimal loss function across multiple non-linear variables, thereby allowing the reverse estimation of each variable's contribution within AI-HI collaboration (Mohd Selamat et al., 2020). Compared with traditional multiple regression analysis, GD offers the following advantages:

3.1.2.1 High-dimensional data processing

By presetting iteration counts and learning rates, the algorithm dynamically updates parameters until partial derivatives converge to a minimum, effectively improving model prediction accuracy and robustness.

3.1.2.2 Innovative variable handling

Instead of treating confounding variables—such as the debated moderating/mediating role of PP—as traditional structural equation modeling does, we process them as specific parameters, avoiding methodological disputes (cf. Zhang et al., 2019; Wang et al., 2022).

3.1.2.3 Feature interpretation advantages

Leveraging machine learning capabilities, we quantitatively evaluate the relative importance of different dimensions of professional skills.

To address RQ1 and RQ2, we construct a GD function (see Equation 5) to model the impact of professional skills on knowledge conversion. Here, $J(\beta)$ represents the loss function, β1∼β8 are the parameter estimates for each indicator, and β0 represents the intercept. IL (Innovation Leadership), RB (Relationship Building), TU (Tolerance for Uncertainty), and PO (Passion and Optimism) are related to soft skills. CB (cyber behavior), DA (data analysis), and FL (foreign language) are related to hard skills.

$$J(\beta) = \frac{1}{2n}\sum_{i=1}^{n}(\beta_0 + \beta_1 \cdot IL_i + \beta_2 \cdot RB_i + \beta_3 \cdot TU_i + \beta_4 \cdot PO_1 + \beta_3 \cdot CB_1 + \beta_6 \cdot DA_1 + \beta_7 \cdot FL_i - KC_1)^2 \quad (5)$$

Furthermore, we construct a centrality-based function (see Equation 6) to capture the impact of social network position on collaborative innovation effectiveness, thereby addressing RQ3.

$$NC(i) = \frac{\deg(i)}{n-1} \quad (6)$$

Here, $NC(i)$ represents the relative centrality of sample iii, and $\deg(i)$ represents the absolute centrality of sample i. Subsequently, the GD method is used to construct the Equation 7 for the impact of centrality on dynamic knowledge conversion, reflecting the visibility of individual identity within the platform.

$$J(\beta_0, \beta_1) = \frac{1}{2n} \sum_{i=1}^{n} (\beta_0 + \beta_1 \cdot NC_i - KC_i)^2 \tag{7}$$

To further address RQ3, we construct a GD function (see Equation 8) to model the influence of centrality, hard skills, and soft skills on knowledge conversion, thereby capturing how network position amplifies or suppresses the effectiveness of professional skills.

$$J(\beta) = \frac{1}{2n} \sum_{i=1}^{n} (\beta_0 + \beta_1 \cdot NC_i + \cdots + \beta_8 \cdot FL_i - KC_1)^2 \tag{8}$$

To address RQ4, we construct Equation 9 to represent the subject-object matching problem in digital identity. In this model, *PP* denotes proactive personality.

$$J(\beta) = \frac{1}{2n} \sum_{i=1}^{n} (\beta_0 + \beta_1 \cdot NC_i + \cdots + \beta_8 \cdot FL_i - KC_1)^2 \tag{9}$$

Equations 5, 7–9 are loss functions concerning regression relationships, and their parameters need to be iteratively updated using the GD method (see Equation 10). Here, α is the learning rate, and $\frac{\partial_{J}\beta)}{\partial_\beta}$ is the partial derivative of the loss function with respect to the regression coefficients.

$$\beta := \beta - \alpha \frac{\partial_{J(\beta)}}{\partial_\beta} \tag{10}$$

To address the algorithm's limitations, we implement a three-tier optimization strategy:

3.1.2.4 Sensitivity to initial values

Running the model algorithm at least 100 times to compute the standard deviation and mean of coefficients.

3.1.2.5 Overfitting control

Employing the *Early Stopping* method to prevent overfitting by testing different parameter settings.

3.1.2.6 Parameter tuning

Utilizing Goodfellow's (2016) dynamic adjustment framework to balance model convergence speed and training efficiency.

In summary, the constructed models are closely aligned with the research questions, with the weights of each dimension reflecting the identification of individual roles in AI-HI collaboration. Variations in these weights indicate changes in digital identity construction under the influence of confounding variables.

3.1.3 Personalized digital identity matrix based on residual-macthing analysis

To further address research questions RQ1 through RQ4, we first identify AI-resilient skills using Equation 5 through Equation 10. We then examine how the interaction among individual motivation, structural position, and resilient skills leads to inconsistencies between system-assigned roles and self-perceived identities. This divergence is analyzed in terms of its impact on AI-HI role allocation.

Based on this premise, we construct an identity recognition framework grounded in residual analysis and matching calculation. The specific steps are as follows:

First, we establish the residuals of knowledge innovation effectiveness (see Equation 11).

$$\text{Residual } = KC_{\text{actual}} - KC_{\text{pred}} \tag{11}$$

Among them, KC_{actual} represents actual effectiveness (related to Equation 4), KC_{pred} representing predictive effectiveness (Equation 12), $\sum \beta_n \cdot$ Skills $_n$ represents key skills.

$$KC_{\text{pred}} = 0.04NC + 0.05PP + \sum \beta_n \cdot \text{ Skills }_n \tag{12}$$

Then, we used PP as a moderator to build the NC-Skills-PP match calculation (Equation 13):

$$\text{Mathscore } = \frac{(NC \times \sum \beta_n \cdot \text{ Skills }_n) \times (1+\alpha \times PP) - \min(\text{ Matchscore })}{\max(\text{ Matchscore }) - \min(\text{ Matchscore })} \times 2 \tag{13}$$

By applying GD, we integrate residual analysis and matching calculation into the theoretical framework of digital identity construction. The difference between predicted and actual identity expression KC_{actual} termed identity matching deviation. Residual serves as the basis for subsequent classification in the personalized collaboration matrix.

Based on the combination of residuals and matching scores, each sample is categorized into one of three types, as shown in Table 1.

TABLE 1 Interpretation of digital identity profiles based on the combination of residuals and matching.

Category	Classification logic	Theoretical interpretation
Low performers	High residual + Low matching	Indicates that both the system and the individual fail to recognize the role, leading to weak identity activation or alignment
Peripheral experts	High matching + High residual	Possess relevant skills but hold a marginal structural position; identity expression is misaligned with system recognition
Core innovators	High matching + Low residual	Demonstrates high triadic alignment, with strong knowledge articulation and consistent platform-based identity recognition

3.2 Explanation of variables

3.2.1 Network centrality, proactive personality and professional skill

This study employs nomination surveys and interviews to collect data on NC. The scale for NC is based on the work of Soda

10.3389/fpsyg.2025.1652130

and Zaheer (2012), using the nomination method and presenting the data in matrix form, as referenced in the study by Cangialosi et al. (2021).

The measurement of PP is based on the scale designed by Seibert et al. (2001). According to previous experimental records, the consistency coefficient between the items of the scale is 0.89. This study conducted a convergence test on the scale design, selecting 8 questions with a 6-point attitude scale.

Cyber Behavior (CB), Data Analysis (DA), and Foreign Languages (FL) represent the critical impact of hard skills on innovation. Based on Hendarman and Cantner (2018), this study collects measurable data to represent hard skills. The measurement of soft skills follows the scale designed by Hendarman and Cantner (2018) and gathers data through questionnaires. Item 1 measures Innovation Leadership (IL), Item 2 measures Relationship Building (RB), Item 3 measures Tolerance for Uncertainty (TU), and Item 4 measures Passion and Optimism (PO). While referencing Hendarman's scale, this study makes adjustments, setting the attitude scale at 6 levels. A detailed structure of the questionnaire design can be found in Supplementary Table S1.

3.2.2 Quantification of cognition in knowledge conversion

As summarized earlier, knowledge conversion is a significant manifestation of knowledge innovation. Therefore, this study employs the After Action Review (AAR) method to conduct a feedback-based evaluation of the knowledge conversion process. The method is applied with reference to Keiser and Arthur (2021), combining Bloom's cognitive model to perform AAR analysis and extract results from the knowledge conversion process. For detailed reference, see Supplementary Table S2.

3.3 Survey participants

The experiment of this study was conducted between 2022 and 2024, targeting companies and employees within the Xiaoguishan Financial Industry Park in Wuhan. The enterprises in this industrial park predominantly belong to symbiotic industries, including project consulting, digital technology, finance, and cultural investment. To align with the digital transformation of industries, the park's enterprises engaged in digital transformation-related work during the pandemic period. The survey involved a total of 455 participants, with 418 valid samples collected. These participants were distributed across 7 community networks, as detailed in Table 2.

It should be noted that most of the participants are interns from Wuhan University of Business and Technology, who underwent similar training and cognitive internships prior to their placement (the average value of forward transfer is 1.00, indicating the coverage of the preliminary training). The majority of participants volunteered to take part, and our survey activities were governed by the regulations of the Hubei Provincial Department of Education project (2021GA078).

TABLE 2 Demographics of the survey respondents.

Variables	AVG	S.D	Industrial distribution	Percentage
Gender	1.67	0.47	Financial	25%
Age	2.06	0.32	Cultural tourism	25%
Education (Edu.)	1.90	0.41	Digitization	13%
Job description (JD)	1.72	1.35	Project consultancy	17%
Forward transfer (FT)	1.00	0.05	Others	20%

N = 418. Gender: 1 = male, 2 = female. Age (years old): 2 = 19–29, 3 = 30–39. Education: 1 = less than bachelor's degree, 2 = bachelor's degree. Job Description: 1 = General Employee, 2 = Departmental Management. Forward Transfer: 1 = had relevant experience and knowledge base.

4 Data analysis

4.1 Data preprocessing

4.1.1 Reliability and validity testing

Table 3 includes the KMO values, total variance explained, AVE, and CR values for the variables in the hypothetical research model. The KMO values indicate a strong correlation among the indicators within the four latent variables of the hypothetical model. The total variance explained (single component) is above 50%, demonstrating that the indicators sufficiently explain the latent variables. AVE values are all above 0.5, and CR values exceed 0.6, indicating that the convergent validity meets the requirements, further validating the scientific rigor of the selected indicators.

4.1.2 Correlation analysis

Based on the heat map of variable correlations (see Figure 3), it is evident that all variables, except for demographic variables, exhibit a certain degree of positive correlation. According to Gelman and Hill (2006, pp. 20–21), the entropy weight method and GD can eliminate the influence of demographic variables on the hypothetical research.

TABLE 3 Reliability and validity of latent variables.

Variables	Items	KMO	TVE	AVE	CR
Knowledge conversion (KC)	6	0.83	77.37 %	0.77	0.93
Proactive personality (PP)	8	0.89	53.70 %	0.53	0.90
Soft skill (SS)	4	0.77	63.00%	0.63	0.87
Hard skill (HS)	3	0.64	65.20 %	0.65	0.85

The analysis results are derived from SPSS and EXCEL. KMO > 0.6; TVE (%) > 50; AVE > 0.5; CR > 0.6.

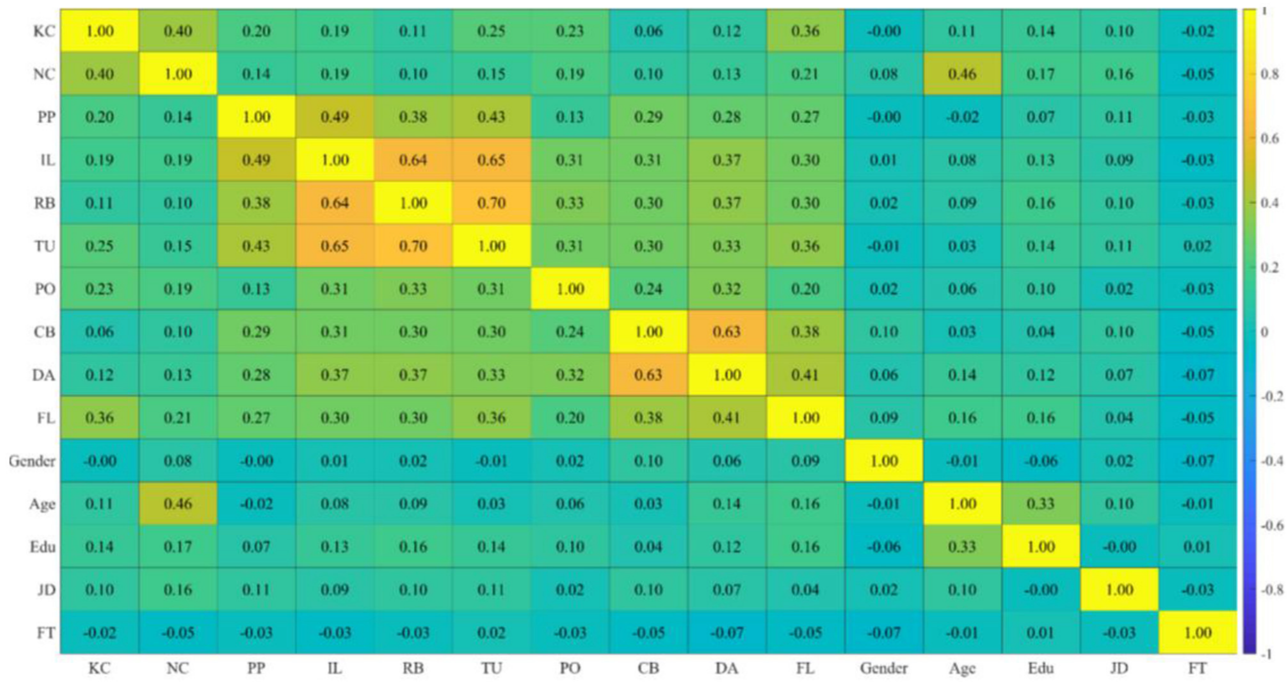


FIGURE 3
Heat map of correlation. The color gradient represents the strength of Pearson correlation (ranging from −1 to +1, from blue to yellow). For variable abbreviations, see Tables 2, 4; NC denotes network centrality, and PP refers to proactive personality.

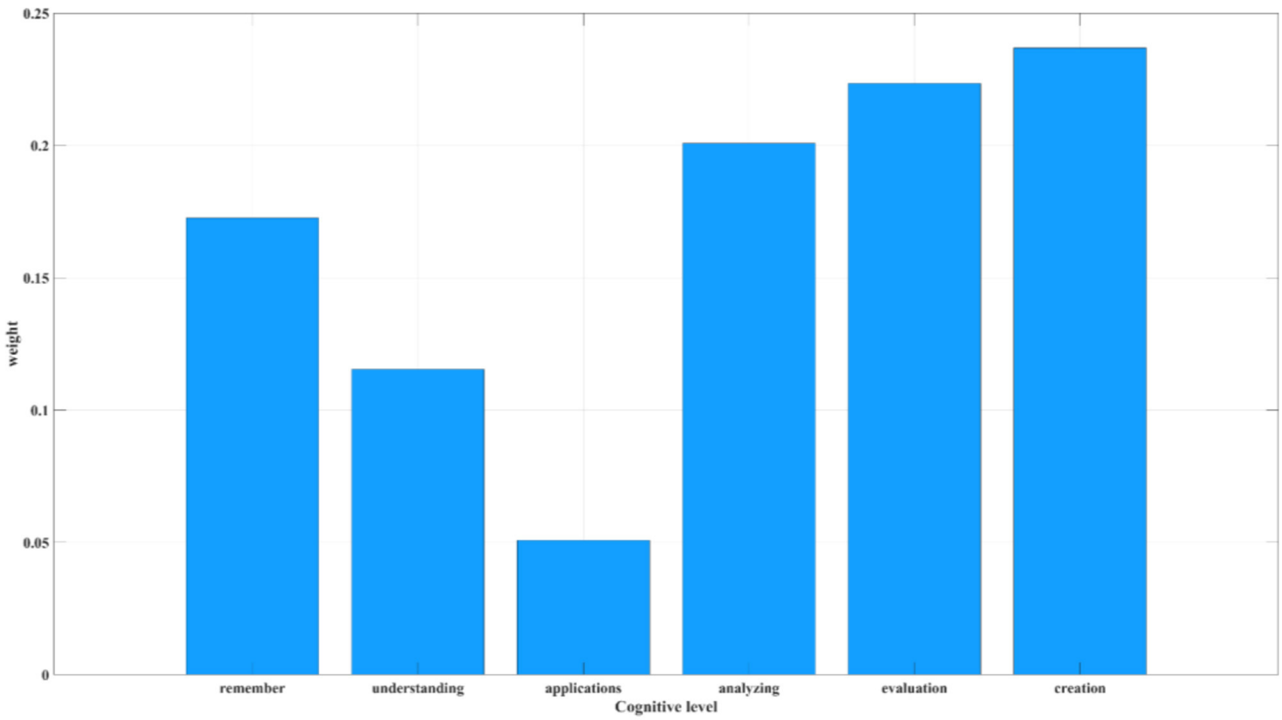


FIGURE 4
Weighted contributions of individual cognitive categories in AI-HI collaboration (Based on entropy weight method). The six cognitive levels are classified according to Bloom's taxonomy. The weight values represent each level's contribution to AI-HI knowledge conversion. Higher-weighted levels—creation, evaluation, and analysis—correspond to higher-order cognition. These results indicate that AI-HI collaboration relies more heavily on complex cognitive processes.

4.2 Quantitative cognitive analysis in the AI-HI collaboration process

Based on Equation 1 through Equation 3, the entropy weighting method was used to quantify individual contributions at different cognitive levels within AI-HI collaboration (see Figure 4). The analysis revealed the following weighted contributions: creation (0.2369) > evaluation (0.2235) > analysis (0.2008) > remembering (0.1726) > understanding (0.1154) > applying (0.0508). These results indicate that, compared to lower-order cognitive processes (remembering, understanding, and applying), higher-order cognition—represented by creation, evaluation, and analysis—plays a more significant role in AI-HI collaborative innovation. This finding reflects AI's substitution effect on lower-order cognition and its enhancement of higher-order cognitive functions (Hu et al., 2021; Asrifan et al., 2024; Jia et al., 2024; Ritala et al., 2024).

Overall, this validates the model's rationality in quantifying AI-HI collaboration effectiveness. Finally, according to Equation 4, the weighted values across cognitive levels can be integrated to predict the AI-HI collaborative innovation effectiveness, represented by the KC value.

4.3 Gradient regression analysis results

Based on responses to RQ1 and RQ2, this study employed GD to model the relationship between professional skills and collaborative innovation effectiveness (Equation 5). The model was iterated 100 times to reduce sensitivity to initial values and improve prediction stability. The results are presented in Table 4.

Regarding skill distribution, among soft skills, IL, TU, and PO, along with the hard skill of FL, showed significant positive effects on AI-HI knowledge conversion effectiveness (KC). Notably, PO ($\beta = 0.03^{**}$) and FL ($\beta = 0.07^{**}$) stood out prominently.

These findings reveal: (1) FL plays a critical role in overcoming knowledge barriers and facilitating knowledge flow, serving as a key factor in AI-HI cross-boundary collaboration; (2) PO functions as a stabilizing mechanism for individual identity, significantly enhancing individuals' capacity to engage with AI-HI collaborative content.

These results support RQ1 by identifying PO and FL as the most AI-resistant professional skills, capable of consistently expressing individual identity and knowledge value.

Further, by integrating Equations 6–9, this study examined the effects of structural position (NC) as a dimension of digital identity recognition, and individual intrinsic motivation (PP) on AI-HI collaborative innovation effectiveness. The findings are as follows:

1. The effect of NC on KC was 0.07^{**}, indicating that structural position characteristics positively contribute to AI-HI collaboration. This result supports the first part of RQ3, confirming that social network position has an empowering effect on digital identity recognition.
2. Under the influence of NC (detailed results of the regression coefficients are reported in Supplementary Table S7), TU (0.03^{**}) and FL (0.05^{**}) showed significant positive impacts on KC. This suggests that, with structural position empowerment, individuals' knowledge articulation is further strengthened, addressing RQ2 regarding the constructive role of the triadic synergy in digital identity formation.
3. Within the triadic synergy framework, regression coefficients for PP, NC, FL, and TU were 0.05^{**}, 0.04^{**}, 0.05^{**}, and 0.03^{*} respectively (detailed results of regression coefficients are shown in Supplementary Table S8). This finding responds to the combined research questions RQ2 and RQ3, illustrating the logic by which the triadic synergy influences individual digital identity recognition.

Note: Details regarding model parameter settings, gradient optimization, and overfitting tests are provided in the Supplementary Tables S3–S6 and Supplementary Figures S1–S4. These materials are intended to validate the model's stability and do not affect the presentation of the main data analysis results.

4.4 Parameter settings for confounding variables and skill priority optimization

To further strengthen the explanatory power of structural position and individual intrinsic motivation on knowledge articulation, this study employed a fixed confounding variable parameter combined with gradient-based re-optimization to

TABLE 4 Contributions of professional skills in the AI-HI collaboration process.

Variable	Definition	Coefficient	Standard error	T value	P value
IL	Innovation leadership	0.02	0.01	1.54	0.12
RB	Relationship building	-0.03^{*}	0.01	−2.19	0.03
TU	Tolerance for uncertainty	0.04^{**}	0.01	3.14	0.00
PO	Passion and optimism	0.03^{**}	0.01	4.63	0.00
CB	Cyber behavior	−0.02	0.01	−1.80	0.07
DA	Data analysis	0.00	0.01	0.39	0.70
FL	Foreign language	0.07^{**}	0.01	7.32	0.00

The analysis results are derived from MATLAB. A smaller standard error indicates a more precise estimate of the regression coefficients; the T value is used to test whether the regression coefficients are significantly different from zero; $^{**}p < 0.01$, $^{*}p < 0.05$.

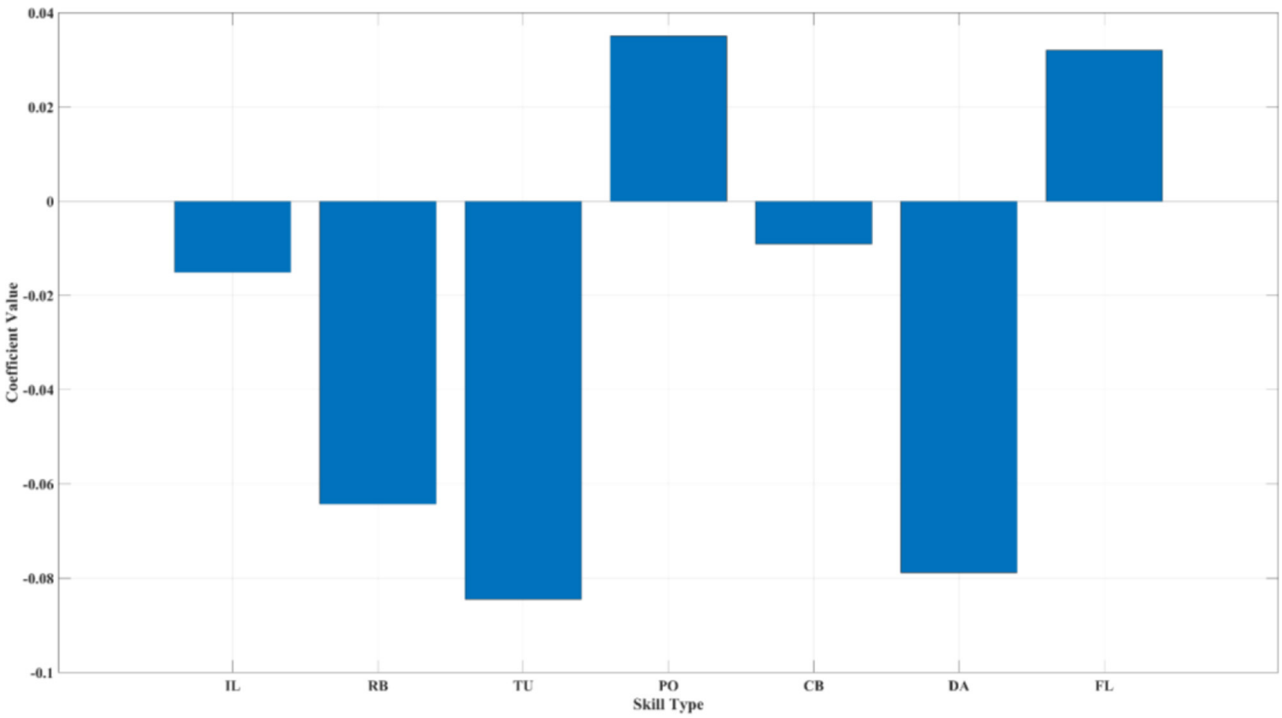


FIGURE 5
Relative contribution ranking of professional skills based on gradient optimization (500 iterations). The Y-axis represents the gradient descent regression coefficients. Both *FL* and *PO* retain significant advantages even under the control of confounding variables *NC* and *PP*, indicating that they are AI-resistant skills. Detailed results of regression coefficients are shown in Supplementary Table S9.

reassess the contribution of professional skills. Specifically, the parameters for NC (0.04**) and PP (0.05**) were fixed. Using gradient regression, the triadic synergy model in Equation 9 was iterated 500 times.

The analysis results (see Figure 5) indicate that, under the conditions of PP and NC, PO (0.04**) and FL (0.03**) emerge as the most "AI-resistant" skills in AI-HI collaboration. This finding further deepens the response to RQ1, demonstrating that "AI-resistant" skills continue to have a stable driving effect on AI-HI collaborative innovation effectiveness even when controlling for structural position and individual motivation. This reflects that truly AI-resistant skills possess strong digital identity expression.

4.5 Decision strategy matrix: classification based on digital identity recognition

In response to RQ4, this study constructs a personal digital identity recognition matrix by combining residuals and matching degrees. The recognition matrix categorizes roles within AI-HI collaboration. Specifically, through parameter settings of confounding variables and skill priority optimization, PO (0.04**) and FL (0.03**) are identified as important indicators of individual knowledge articulation. Then, based on Equations 11–13, the distribution of sample sizes, strategy classifications, and adaptation recommendations for digital identity recognition are obtained (see Figure 6, Table 5).

Table 5 details the proportions and response strategies for Core Innovators, Marginal Experts, and Low Performers. These results demonstrate that digital identities under the triadic synergy framework can be quantitatively identified and dynamically adjusted, providing effective recommendations for role allocation in AI-HI collaboration.

5 Discussion, conclusion, and outlook

5.1 Summary of key findings

This study investigates digital identity recognition under AI-HI collaborative innovation by constructing a triadic mechanism involving knowledge articulation, structural position, and personal motivation. Compared to previous research, it offers key findings and extensions in the following aspects:

First, we redefine "AI-resistant" professional skills, breaking through the traditional binary perspective of soft vs. hard skills. Prior studies primarily focused on the structural differences between soft and hard skills (Hendarman and Cantner, 2018; Alekseeva et al., 2021; Fletcher and Thornton, 2023). Our model, incorporating confounding variables, reveals that FL and PO exhibit more stability in AI-HI collaboration. This finding acknowledges the unique roles of both hard and soft skills in AI-HI collaboration and highlights that highly expressive and system-recognized skills are critical in constructing digital identity. Furthermore, while Alekseeva et al. emphasize the key roles of DA and CB in the digital economy, our findings suggest these

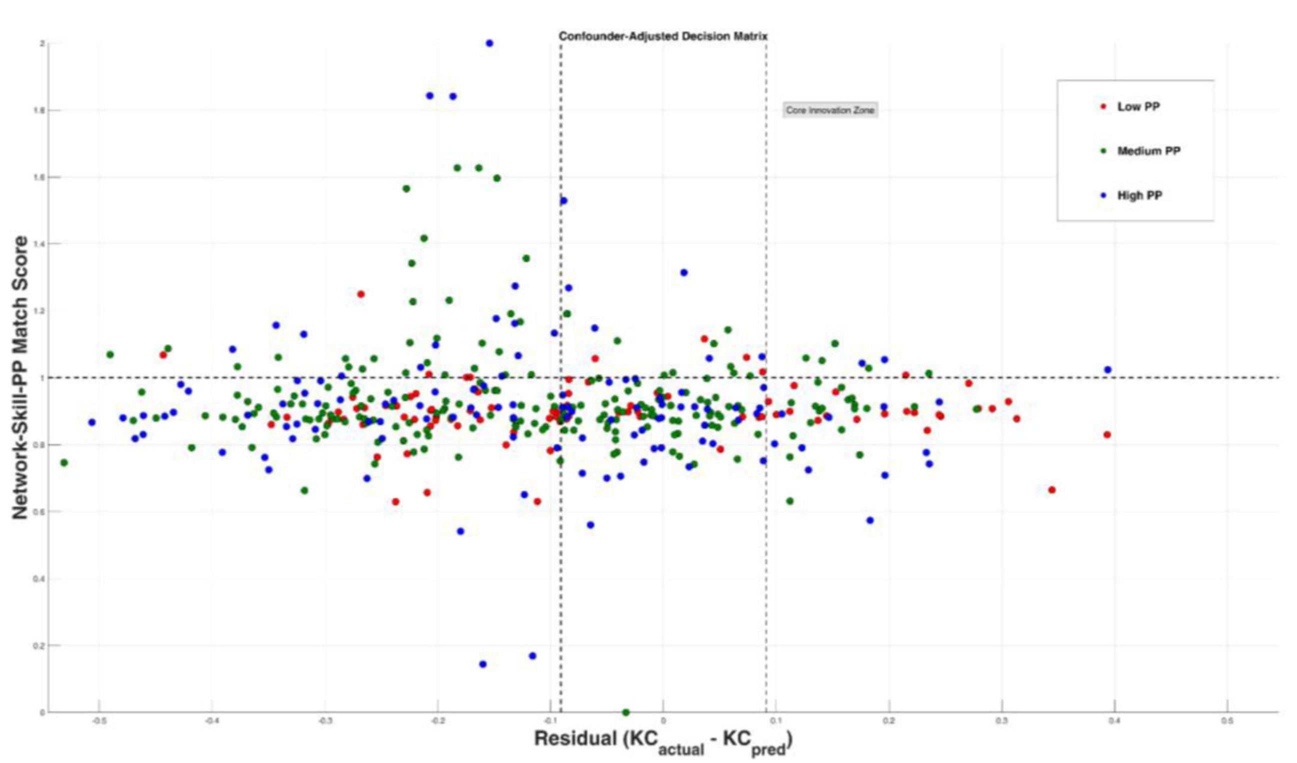


FIGURE 6
Sample size distribution based on residuals and matching degree. The X-axis represents the degree of identity matching, while the Y-axis indicates the residuals of innovation effectiveness. Marginal experts are primarily located in the "high residual + medium matching" quadrant, revealing a structural misalignment in identity recognition.

TABLE 5 Digital identity strategy types and adaptation recommendations.

Strategy type	Low PP group (15 %)	Mid-PP group (45%)	High PP group (40 %)
Core innovators	0	0.01	0.99
Marginal experts	0.12	0.23	0.65
Low-efficiency individuals	0.82	0.15	0.03
Residuals\Match	**Low match (MatchScore < 0.8)**	**Medium match (0.8 ≤ MatchScore ≤ 1.2)**	**High match (MatchScore > 1.2)**
High residual (Residual > +0.3)	Marginal experts: AI-enhanced breakthroughs	Core innovators: strategic ownership	Core innovators: strategic ownership
Medium residual (−0.3 ≤ Residual ≤ +0.3)	Routine optimisation	Progressivity	Structured tasks
Low residuals (Residual < −0.3)	Low-efficiency individuals: redeployment/elimination	Standardized training	Rotation activation

skills are increasingly supplanted by AI, whereas FL and PO show stronger resistance to replacement. This provides empirical support for research on AI-HI complementarity.

It is important to clarify that FL as an AI-resistant skill may be easily misunderstood—after all, LLMs are widely recognized as one of AI's core strengths. However, FL in this study is not limited to linguistic proficiency *per se*. Instead, it represents a deeper capability: the power to integrate knowledge across boundaries (Nonaka and Yamaguchi, 2022). In contexts where AI plays a dominant role but lacks contextual adaptability and cultural discernment, FL is more likely to function as a technological tool for identity expression and bridging, enabling individuals to serve as connectors in human-AI collaboration.

Second, we propose a collaborative construction logic of digital identity, critiquing prior studies' reliance on "system-label" identity assignment. Typically, role recognition and assignment in human–AI collaboration are attributed to system-level configurations (e.g., Alowais et al., 2023; Bharatha et al., 2024). However, such studies often overlook the interactive dynamics among self-driven motivation, collective dynamics, and cognitive performance. Our study argues that individual digital identity emerges from the synergy of internal motivation, structural position, and knowledge articulation. Moreover, by combining residual and matching degree analyses, we concretize the discrepancy between system recognition and individual identity, empirically challenging the simplistic "system-label" model.

Third, this study reveals a bidirectional "empowering/inhibiting" effect of structural position on individual knowledge articulation, offering a contextualized extension of the "structural empowerment logic" within existing social capital theory. In contrast to prior studies (e.g., Cangialosi et al., 2021), our findings suggest that in hybrid human–AI social networks, NC does not uniformly activate knowledge articulation. Its amplifying effect depends on skill fit and personal motivation. The identification of "marginal experts" in Figure 6 and Table 5 illustrates that, even individuals with high-level skills may find their identity expression suppressed—particularly when they occupy structurally marginal positions and lack proactive personal drivers.

Additionally, RB also contributes to the expression of structural position. However, in our model, it exhibits a negative effect ($\beta = -0.03$). We speculate that in highly collaborative and digitized task environments, an overreliance on RB may lead to reduced communication efficiency, blurred role boundaries, and a tendency to avoid conflict. This finding aligns with previous research suggesting that unstructured social interactions in virtual teams can undermine collaborative performance (Morrison-Smith and Ruiz, 2020; Caldwell et al., 2022).

Fourth, establishing an intervention mechanism for digital identity recognition addresses the current limitations in quantifying AI–HI collaborative innovation effectiveness and the challenges posed by the "black box" nature of evaluations. Existing studies largely rely on self-reported questionnaires to measure collaborative innovation outcomes (Chowdhury et al., 2022; Pham et al., 2024), which do not fully capture the cognitive dynamics and human–AI role distinctions inherent in AI–HI collaboration. This study employs the entropy weight method to develop a more objective evaluation framework for human–AI collaborative effectiveness. Furthermore, through a residual–matching analysis matrix, it identifies three distinct groups—core innovators, marginal experts, and low performers. This approach not only enhances the objectivity and explanatory power of AI-driven performance assessments but also offers practical insights for the design of AI–HI symbiotic mechanisms and human–AI role allocation.

5.2 Theoretical contributions

First, this study reconstructs the AI-HI collaboration theoretical framework from the perspective of digital identity. It breaks the traditional binary paradigm of technological determinism and capability determinism, proposing a triadic synergy model of "Motivation-Structure- Articulation," embedding digital identity construction into the AI-HI symbiotic ecosystem.

Second, methodological innovations are made by integrating entropy weighting, gradient descent, and residual-matching techniques. This not only improves self-reported measurement approaches but also innovatively quantifies digital identity recognition in AI-HI collaborative scenarios.

Finally, the study redefines the skill substitution effect. It argues that language proficiency and emotional intrinsic motivation serve as more robust AI-resistant skills in AI-HI collaboration than procedural skills such as data analysis. This shift helps advance the research paradigm from "technical adaptation" toward "identity expression" in professional skills studies.

5.3 Management insight

First, based on the residual-matching decision matrix, a personalized talent identification and allocation mechanism can be constructed to enable precise interventions for "core innovators," "marginal experts," and "low performers." Core innovators should be granted greater autonomy; low performers should undergo standardized training, job rotation, or optimization; and marginal experts should be supported through a structured mentorship system.

Second, in designing AI-driven collaborative effectiveness, structural adjustments of AI-HI social networks are necessary, especially by providing platforms and collaboration opportunities for high-skill but low-position marginal experts. To enhance the structural recognition of marginal experts, cross-departmental task forces should be established to increase their involvement in key roles. Additionally, network-building initiatives can be implemented to facilitate their integration into core collaboration circles.

Finally, combining residual-matching strategies supports sustainable talent-resource allocation. In AI-HI skill training, priority should be given to cultivating AI-resistant skills

5.4 Shortcomings and outlook

First, the assessment of soft skills remains subjective. The evaluation methods and criteria for soft skills in this study are primarily based on questionnaires, which are evidently overly subjective. Therefore, future research should adopt more objective approaches to quantify soft skills, such as technology-driven soft skill assessments (Altomari et al., 2023).

Secondly, the triadic synergy model effectively explains digital identity construction in AI-HI collaboration. However, variables such as AI model performance and task scenarios require further refinement. Moreover, this study primarily employs quantitative analysis and lacks the integration of qualitative data (e.g., interviews), which is essential for enhancing the explanatory power of the model. At the same time, future research should consider conducting group-based comparative tests on the residual–matching approach to improve the model's robustness and rigor.

Third, the cross-sectional nature of this study limits a comprehensive understanding of AI technology development and iteration over time. Moreover, the impact of Explainable Artificial Intelligence (XAI) on knowledge innovation exhibits distinct phase-specific characteristics (Mancuso et al., 2025). Therefore, future research should incorporate longitudinal studies to observe temporal changes in relevant variables.

Fourth, this study did not quantitatively assess the technical characteristics of AI systems, although existing research highlights their close relationship with digital identity construction and task adaptation (Shneiderman, 2022; Huang and Rust, 2024). The diversity of AI-HI collaborative systems poses a significant challenge for current research. Future studies should aim to develop appropriate and generalizable methods to quantify AI system features, enabling a more comprehensive characterization of AI-HI collaboration mechanisms.

Finally, while this study's sample size is substantial, it is limited by the convenience of data collection, resulting in a relatively homogeneous respondent background and role distribution. Although participants generally received relevant training, many occupied peripheral positions within social network structures and had insufficient understanding of work processes. These factors inevitably impact the knowledge expression system and collaboration effectiveness. Future research should consider cross-validating findings across multiple industries and among formal employees to further strengthen the robustness and generalizability of the conclusions.

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 humans were approved by the Research Office of Wuhan Technology and Business University. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

WJ: Conceptualization, Data curation, Writing – original draft, Writing – review & editing, Formal analysis, Investigation, Methodology, Visualization. YL: Conceptualization, Formal analysis, Validation, Writing – review & editing. XH: Methodology, Writing – review & editing. DM: Methodology, Writing – review & editing.

Funding

The author(s) declare that no financial support was received for the research and/or publication of this article.

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.

Generative AI statement

The author(s) declare that Gen AI was used in the creation of this manuscript. During the preparation of this manuscript, the author(s) utilized the following tools for text translation assistance: DeepL (web version), and Youdao Translate (web version). DeepL and Youdao Translate have been embedded with AI tools. The author(s) have thoroughly reviewed and edited the output and take full responsibility for the content of this publication.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2025.1652130/full#supplementary-material

References

Alekseeva, L., Azar, J., Giné, M., Samila, S., and Taska, B. (2021). The demand for AI skills in the labor market. *Labour Econ.* 71:102002. doi: 10.1016/j.labeco.2021.102002

Alowais, S. A., Alghamdi, S. S., Alsuhebany, N., Alqahtani, T., Alshaya, A. I., Almohareb, S. N., et al. (2023). Revolutionizing healthcare: the role of artificial intelligence in clinical practice. *BMC Med. Educ.* 23:689. doi: 10.1186/s12909-023-04698-z

Altomari, L., Altomari, N., and Iazzolino, G. (2023). Gamification and soft skills assessment in the development of a serious game: design and feasibility pilot study. *JMIR Serious Games* 11:e45436. doi: 10.2196/45436

Andrea, A., Aaron De, S., Meredith, L., and Dan, R. (2021). *What employees are saying about the future of remote work.* McKinsey and Company. Available online at: https://www.mckinsey.com/capabilities/people-and-organizational-performance/our-insights/what-employees-are-saying-about-the-future-of-remote-work (Accessed May 10, 2024).

Asrifan, A., Rusmayadi, R., Hasmawaty, H., Nirwana, N., Sivaraja, S., and Dewi, A. C. (2024). "Synergizing human and machine: innovations in knowledge discovery and management," in *Advances in Environmental Engineering and Green Technologies*, eds. R. Kumar, A. B. Abdul Hamid, N. I. Binti Ya'akub, T. E. Nyamasvisva, and R. K. Tiwari (Makassar: IGI Global), 249–282. doi: 10.4018/979-8-3693-9879-1.ch010

Bahoo, S., Cucculelli, M., and Qamar, D. (2023). Artificial intelligence and corporate innovation: a review and research agenda. *Technol. Forecast. Soc. Change* 188:122264. doi: 10.1016/j.techfore.2022.122264

Bankins, S., Ocampo, A. C., Marrone, M., Restubog, S. L. D., and Woo, S. E. (2024). A multilevel review of artificial intelligence in organizations: implications for organizational behavior research and practice. *J. Organ. Behav.* 45, 159–182. doi: 10.1002/job.2735

Bharatha, A., Ojeh, N., Fazle Rabbi, A. M., Campbell, M., Krishnamurthy, K., Layne-Yarde, R., et al. (2024). Comparing the performance of ChatGPT-4 and medical students on MCQs at varied levels of bloom's taxonomy. *Adv. Med. Educ. Pract.* 15, 393–400. doi: 10.2147/AMEP.S457408

Burtch, G., Lee, D., and Chen, Z. (2023). The consequences of generative AI for UGC and online community engagement. *SSRN Electron. J.* doi: 10.2139/ssrn.4521754

CAICT (2024). *Artificial Intelligence Development Report (2024 Edition).* Beijing, China: China Academy of Information and Communications Technology. Available online at: http://www.caict.ac.cn/kxyj/qwfb/bps/202412/P020241210548865982463.pdf

Caldwell, S., Sweetser, P., O'Donnell, N., Knight, M. J., Aitchison, M., Gedeon, T., et al. (2022). An agile new research framework for hybrid human-AI teaming: trust, transparency, and transferability. *ACM Trans. Interact. Intell. Syst.* 12, 1–36. doi: 10.1145/3514257

Cangialosi, N., Odoardi, C., Battistelli, A., and Baldaccini, A. (2021). The social side of innovation: when and why advice network centrality promotes innovative work behaviours. *Creat. Innov. Manag.* 30, 336–347. doi: 10.1111/caim.12434

Carlisle, S., Ivanov, S., and Dijkmans, C. (2023). The digital skills divide: evidence from the European tourism industry. *J. Tour. Futur.* 9, 240–266. doi: 10.1108/JTF-07-2020-0114

Chen, P. (2021). Effects of the entropy weight on TOPSIS. *Expert Syst. Appl.* 168:114186. doi: 10.1016/j.eswa.2020.114186

Chowdhury, S., Budhwar, P., Dey, P. K., Joel-Edgar, S., and Abadie, A. (2022). AI-employee collaboration and business performance: integrating knowledge-based view, socio-technical systems and organisational socialisation framework. *J. Bus. Res.* 144, 31–49. doi: 10.1016/j.jbusres.2022.01.069

Chu, J., Wu, R., and Li, Z. (2023). Study of the constitution and mechanism of the innovation ecosystem empowered by big data and intellectual technology. *Inf. Stud. Appl.* 46, 1–8. Available online at: https://qikan.cqvip.com/Qikan/Article/Detail?id=7109126306

Chuang, S. (2024). Indispensable skills for human employees in the age of robots and AI. *Eur. J. Train. Dev.* 48, 179–195. doi: 10.1108/EJTD-06-2022-0062

Clark, J., Perrault, R., and Brynjolfsson, E. (2024). *AI Index Report 2024.* Stanford, CA: Stanford University Human-Centered Artificial Intelligence. Available online at: https://aiindex.stanford.edu/report/

Escolà-Gascón, Á., and Gallifa, J. (2022). How to measure soft skills in the educational context: psychometric properties of the SKILLS-in-ONE questionnaire. *Stud. Educ. Eval.* 74:101155. doi: 10.1016/j.stueduc.2022.101155

Fletcher, S., and Thornton, K. R. V. (2023). The top 10 soft skills in business today compared to 2012. *Bus. Prof. Commun. Q.* 86, 411–426. doi: 10.1177/23294906221148453

Gelman, A., and Hill, J. (2006). *Data Analysis Using Regression and Multilevel/Hierarchical Models.* 1st Edn. New York, NY: Cambridge University Press. doi: 10.1017/CBO9780511790942

Goodfellow, I., Bengio, Y., and Courville, A. (2016). *Deep Learning.* Cambridge, Mass: The MIT press.

Gray, R. M. (2011). *Entropy and Information Theory.* Boston, MA: Springer US. doi: 10.1007/978-1-4419-7970-4

Hendarman, A. F., and Cantner, U. (2018). Soft skills, hard skills, and individual innovativeness. *Eurasian Bus. Rev.* 8, 139–169. doi: 10.1007/s40821-017-0076-6

Hu, Q., Lu, Y., Pan, Z., Gong, Y., and Yang, Z. (2021). Can AI artifacts influence human cognition? The effects of artificial autonomy in intelligent personal assistants. *Int. J. Inf. Manag.* 56:102250. doi: 10.1016/j.ijinfomgt.2020.102250

Huang, M.-H., and Rust, R. T. (2024). The caring machine: feeling AI for customer care. *J. Mark.* 88, 1–23. doi: 10.1177/00222429231224748

iResearch Group (2024). *2024 China Collaborative Office Platform Industry Research Report*. Shanghai, China: iResearch Group. Available online at: https://report.iresearch.cn/report_pdf.aspx?id=4655 (Accessed May 10, 2024).

Jia, N., Luo, X., Fang, Z., and Liao, C. (2024). When and how artificial intelligence augments employee creativity. *Acad. Manage. J.* 67, 5–32. doi: 10.5465/amj.2022.0426

Jiang, W., Li, D., and Liu, C. (2025). Understanding dimensions of trust in AI through quantitative cognition: Implications for human-AI collaboration. *PLoS ONE* 20:e0326558. doi: 10.1371/journal.pone.0326558

Keiser, N. L., and Arthur, W. (2021). A meta-analysis of the effectiveness of the after-action review (or debrief) and factors that influence its effectiveness. *J. Appl. Psychol.* 106, 1007–1032. doi: 10.1037/apl0000821

Kong, H., Yin, Z., Chon, K., Yuan, Y., and Yu, J. (2024). How does artificial intelligence (AI) enhance hospitality employee innovation? The roles of exploration, AI trust, and proactive personality. *J. Hosp. Mark. Manag.* 33, 261–287. doi: 10.1080/19368623.2023.2258116

Lythreatis, S., Singh, S. K., and El-Kassar, A.-N. (2022). The digital divide: a review and future research agenda. *Technol. Forecast. Soc. Change* 175:121359. doi: 10.1016/j.techfore.2021.121359

Mancuso, I., Petruzzelli, A. M., Panniello, U., and Frattini, F. (2025). The role of explainable artificial intelligence (XAI) in innovation processes: a knowledge management perspective. *Technol. Soc.* 82:102909. doi: 10.1016/j.techsoc.2025.102909

Mohd Selamat, S. A., Prakoonwit, S., and Khan, W. (2020). A review of data mining in knowledge management: applications/findings for transportation of small and medium enterprises. *SN Appl. Sci.* 2:818. doi: 10.1007/s42452-020-2589-3

Morrison-Smith, S., and Ruiz, J. (2020). Challenges and barriers in virtual teams: a literature review. *SN Appl. Sci.* 2:1096. doi: 10.1007/s42452-020-2801-5

Ng, Y.-L. (2022). Exploring the association between use of conversational artificial intelligence and social capital: survey evidence from Hong Kong. *New Media Soc.* 26:146144482210740. doi: 10.1177/14614448221074047

Nonaka, I., and Yamaguchi, I. (2022). "The SECI Model: Knowledge Creation in the Cycle of Subjectivity and Objectivity," in *Management by Eidetic Intuition* (Singapore: Springer Nature Singapore), 125–139. doi: 10.1007/978-981-16-6851-7_10

Okamura, K., and Yamada, S. (2020). Adaptive trust calibration for human-AI collaboration. *PLoS ONE* 15:e0229132. doi: 10.1371/journal.pone.0229132

Pham, Thao, P. T., Van Nguyen, T., Van Nguyen, P., and Ahmed, Z. U. (2024). The pathways to innovative work behavior and job performance: exploring the role of public service motivation, transformational leadership, and person-organization fit in Vietnam's public sector. *J. Open Innov. Technol. Mark. Complex.* 10:100315. doi: 10.1016/j.joitmc.2024.100315

Rai, A. (2020). Explainable AI: from black box to glass box. *J. Acad. Mark. Sci.* 48, 137–141. doi: 10.1007/s11747-019-00710-5

Rienda, L., Andreu, R., and Yang, Y. (2025). Entrepreneurs' proactive personality, network involvement, and international entrepreneurial intention in family SMEs. *Rev. Manag. Sci.* doi: 10.1007/s11846-025-00899-6

Ritala, P., Ruokonen, M., and Ramaul, L. (2024). Transforming boundaries: how does ChatGPT change knowledge work? *J. Bus. Strategy* 45, 214–220. doi: 10.1108/JBS-05-2023-0094

Sedlmeir, J., Smethurst, R., Rieger, A., and Fridgen, G. (2021). Digital identities and verifiable credentials. *Bus. Inf. Syst. Eng.* 63, 603–613. doi: 10.1007/s12599-021-00722-y

Seibert, S. E., Kraimer, M. L., and Crant, J. M. (2001). What do proactive people do? a longitudinal model linking proactive personality and career success. *Pers. Psychol.* 54, 845–874. doi: 10.1111/j.1744-6570.2001.tb00234.x

Sheng, X., Murray, R., Ketron, S. C., and Felix, R. (2024). Effects of third-party observer empathy when viewing interactions between robots and customers: the moderating role of robot eeriness. *Int. J. Hosp. Manag.* 119:103729. doi: 10.1016/j.ijhm.2024.103729

Shneiderman, B. (2022). *Human-Centered AI., 1st Edn.* Oxford University Press: Oxford. doi: 10.1145/3538882.3542790

Singla, A., Sukharevsky, A., Yee, L., Chui, M., and Hall, B. (2025). *The state of AI: How organizations are rewiring to capture value*. McKinsey and Company. Available online at: https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai (Accessed May 5, 2025).

Soda, G., and Zaheer, A. (2012). A network perspective on organizational architecture: performance effects of the interplay of formal and informal organization. *Strateg. Manag. J.* 33, 751–771. doi: 10.1002/smj.1966

Tariq, S., Baruwal Chhetri, M., Nepal, S., and Paris, C. (2025). A 2 C: a modular multi-stage collaborative decision framework for human–AI teams. *Expert Syst. Appl.* 282:127318. doi: 10.1016/j.eswa.2025.127318

Uhl-Bien, M. (2021). Complexity leadership and followership: changed leadership in a changed world. *J. Change Manag.* 21, 144–162. doi: 10.1080/14697017.2021.1917490

Wang, W., Ning, H., Shi, F., Dhelim, S., Zhang, W., and Chen, L. (2022). A survey of hybrid human-artificial intelligence for social computing. *IEEE Trans. Hum.-Mach. Syst.* 52, 468–480. doi: 10.1109/THMS.2021.3131683

Zeng, J., and Yang, J. (2024). English language hegemony: retrospect and prospect. *Humanit. Soc. Sci. Commun.* 11:317. doi: 10.1057/s41599-024-02821-z

Zhang, A., Li, X., and Guo, Y. (2021). Proactive personality and employee creativity: a moderated mediation model of multisource information exchange and LMX. *Front. Psychol.* 12:552581. doi: 10.3389/fpsyg.2021.552581

Zhang, J., Bal, P. M., Akhtar, M. N., Long, L., Zhang, Y., and Ma, Z. (2019). High-performance work system and employee performance: the mediating roles of social exchange and thriving and the moderating effect of employee proactive personality. *Asia Pac. J. Hum. Resour.* 57, 369–395. doi: 10.1111/1744-7941.12199

Zheng, F., Khan, N. A., and Hussain, S. (2020). The COVID 19 pandemic and digital higher education: Exploring the impact of proactive personality on social capital through internet self-efficacy and online interaction quality. *Child. Youth Serv. Rev.* 119:105694. doi: 10.1016/j.childyouth.2020.105694

Zhu, Y., Tian, D., and Yan, F. (2020). Effectiveness of Entropy Weight Method in Decision-Making. *Math. Probl. Eng.* 2020, 1–5. doi: 10.1155/2020/3564835

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 29 August 2025
DOI 10.3389/fpsyg.2025.1454197

OPEN ACCESS

EDITED BY
Xi Chen,
Yunnan University, China

REVIEWED BY
Andrian Haro,
Jakarta State University, Indonesia
Cheng Chen,
Yunnan University, China
Cristobal Rodolfo Guerra-Tamez,
University of Monterrey, Mexico

*CORRESPONDENCE
Jialiang Chen
✉ 532986524@qq.com

RECEIVED 08 November 2024
ACCEPTED 07 August 2025
PUBLISHED 29 August 2025

CITATION
Liu J and Chen J (2025) Chatbot-aided product purchases among Generation Z: the role of personality traits.
Front. Psychol. 16:1454197.
doi: 10.3389/fpsyg.2025.1454197

Chatbot-aided product purchases among Generation Z: the role of personality traits

Jinrong Liu[1] and Jialiang Chen[2]*

[1]Department of Physical Education, University of Shanghai for Science and Technology, Shanghai, China, [2]Shanghai University of Sport, Shanghai, China

Introduction: The rapid integration of machine learning has positioned product recommendation chatbots as essential tools in the e-commerce landscape, shaping how consumers engage and make purchasing decisions. Generation Z, as a tech-savvy and AI-adept demographic, plays a central role in this transformation. While prior studies have examined chatbot-consumer interactions, limited research has explored how both personality traits and information source characteristics jointly influence purchase intentions.

Methods: This study develops an integrative framework to assess how the Big Five personality traits—extraversion, agreeableness, conscientiousness, neuroticism, and openness—and key chatbot features—expertise, interactivity, trustworthiness, and customization—affect Generation Z's willingness to purchase chatbot-recommended products. The moderating role of personal innovativeness is also examined. Data were collected from 480 Generation Z chatbot users in China through an online survey and analyzed using structural equation modeling (SEM), artificial neural networks (ANN), and necessary condition analysis (NCA).

Results: Results indicate that extraversion, agreeableness, openness, expertise, interactivity, and customization significantly influence purchase intention. Moreover, personal innovativeness positively moderates the effect of extraversion on purchase intention.

Discussion: These findings contribute to the literature by bridging personality psychology and human–AI interaction and offer practical insights for enhancing chatbot effectiveness in e-commerce.

KEYWORDS

product recommendation chatbots, Generation Z, Big Five personality traits, source characteristics of information, SEM-ANN-NCA

1 Introduction

With the rapid advancement of artificial intelligence (AI) technologies, e-commerce firms are increasingly integrating conversational agents into their digital touchpoints to deliver highly personalized and interactive marketing experiences (Peng et al., 2023; Wang, 2023). Product-recommendation chatbots utilize natural language processing and machine learning algorithms to curate item assortments tailored to individual consumer preferences, thereby enhancing perceived utility and shopping satisfaction (Jin and Eastin, 2023). Real-world implementations exemplify this trend. For example, Lazada, a leading Southeast Asian platform, has launched LazzieChat, a GPT-powered assistant that provides real-time personalized suggestions. Similarly, Taobao has developed Ali Xiaomi, a conversational interface that promotes user engagement through socially oriented dialogs (Yin and Qiu, 2021). Empirical evidence suggests that such emotionally enriched human–AI interactions positively affect consumer affect and purchasing intentions (Wang C. et al., 2023; Wang L. et al., 2023; Wang X. et al., 2023).

In this context, scholarly attention has increasingly focused on how consumers interact with chatbots. Prior studies have demonstrated that variations in chatbot design—such as language style, anthropomorphic cues, and emoticon use—evoke different consumer responses (Li and Shin, 2023; Li and Wang, 2023; Lu et al., 2024). However, users' psychological profiles, socioeconomic backgrounds, and cultural traits also play a critical role in shaping human–computer interaction (Arpaci et al., 2022). While personality traits have been shown to influence engagement with AI-based systems (Arpaci and Kocadag Unver, 2020), their role in chatbot-mediated product recommendations remains insufficiently understood. Simultaneously, chatbots serve as information sources whose perceived characteristics—such as expertise, trustworthiness, and interactivity—can directly influence consumer evaluations and behavioral intentions (Han, 2021; Shin et al., 2023). Thus, considering both personality traits and information-source attributes may provide deeper insights into decision-making mechanisms in AI-mediated commerce.

Despite these insights, two major research gaps remain. First, prior work has tended to examine personality and information-source factors in isolation, failing to explore how they might interactively shape consumer intention. Second, most studies rely on linear modeling techniques that may not adequately capture the complex, nonlinear relationships among psychological and technological variables. These limitations hinder our understanding of how Generation Z responds to chatbot recommendations in dynamic digital environments.

To address these gaps, this study proposes the following research questions: (1) How do personality traits shape Generation Z consumers' intentions to purchase chatbot-recommended products? (2) How do different chatbot information attributes affect purchase intention? (3) Does personal innovativeness moderate the effect of personality traits on purchase decisions? (4) How do different analytic approaches—linear, nonlinear, and necessity-based—converge or diverge in their interpretations of these relationships?

To answer these questions, this study adopts an integrated methodological framework combining partial least squares structural equation modeling (PLS-SEM), artificial neural networks (ANN), and necessary condition analysis (NCA). This multi-method approach enables us to capture both linear and nonlinear relationships, as well as necessary conditions for specific outcomes, thereby offering a more comprehensive understanding of the mechanisms driving Generation Z's behavioral responses to chatbot product recommendations.

2 Research design

In the initial phase of this study, partial least squares structural equation modeling (PLS-SEM) was employed to examine the linear relationships among latent variables and to validate the hypothesized conceptual model. As a prominent variance-based technique within the SEM family, PLS-SEM integrates features of principal component analysis and multiple regression, offering flexibility in handling complex, multi-path models with relatively small sample sizes and non-normally distributed data (Lew et al., 2020). Given the exploratory nature of the present study and the presence of both reflective and formative constructs, PLS-SEM was selected as the primary tool for testing the theoretical pathways linking the Big Five personality traits, chatbot information-source characteristics, and purchase intention.

To complement the linear perspective of PLS-SEM and explore potential nonlinear patterns in the data, artificial neural network (ANN) analysis was incorporated in the second phase. As a data-driven computational technique, ANN is well-suited to model high-order interactions and nonlinear dependencies without imposing distributional assumptions (Lo et al., 2022). In this study, the latent scores extracted from PLS-SEM were used as input features for a three-layer feedforward ANN comprising an input layer (corresponding to the extracted components), a hidden layer, and a single-node output layer predicting purchase intention. The hidden layer adopted the ReLU activation function, while the output layer utilized a Sigmoid function to produce probabilistic outcomes. The network was trained using the Adam optimizer with a learning rate of 0.01, batch size of 32, and early stopping based on validation loss. Ten-fold cross-validation was performed to ensure generalizability, and model fit was evaluated using mean squared error (MSE) and prediction accuracy. In addition, a permutation-based sensitivity analysis was conducted to derive the relative importance of each predictor, providing a ranking of the most influential factors in driving behavioral intention.

Although the combined use of PLS-SEM and ANN allowed for the identification of both linear and nonlinear associations, these approaches do not evaluate whether certain variables constitute indispensable prerequisites for behavioral outcomes. Therefore, in the third analytical phase, necessary condition analysis (NCA) was implemented to determine whether specific antecedents functioned as non-compensatory constraints for the occurrence of high purchase intention (Richter et al., 2020). NCA operates under the logic of necessity rather than sufficiency: it posits that if a necessary condition is not met, the desired outcome cannot occur, regardless of the levels of other predictors. The analysis was conducted using the CE-FDH (free disposal hull) method in RStudio, and significance testing was performed using 10,000-fold permutation sampling. By identifying minimum thresholds that must be exceeded for the outcome to manifest, NCA provides a complementary diagnostic lens that extends beyond correlational inference.

Taken together, the integration of PLS-SEM, ANN, and NCA constitutes a comprehensive, triangulated analytical framework that captures linear causality, nonlinear complexity, and asymmetrical necessity (see Figure 1). This hybrid approach enables a deeper understanding of the multidimensional mechanisms through which personality traits and information source characteristic jointly shape Generation Z's product-purchase decisions in AI-mediated retail environments. Beyond statistical robustness, this design also aligns with theoretical pluralism by combining hypothesis-driven testing with data-centric exploration and constraint-based reasoning.

3 Conceptual background and hypotheses development

As e-commerce platforms increasingly embed artificial intelligence technologies into their service architecture,

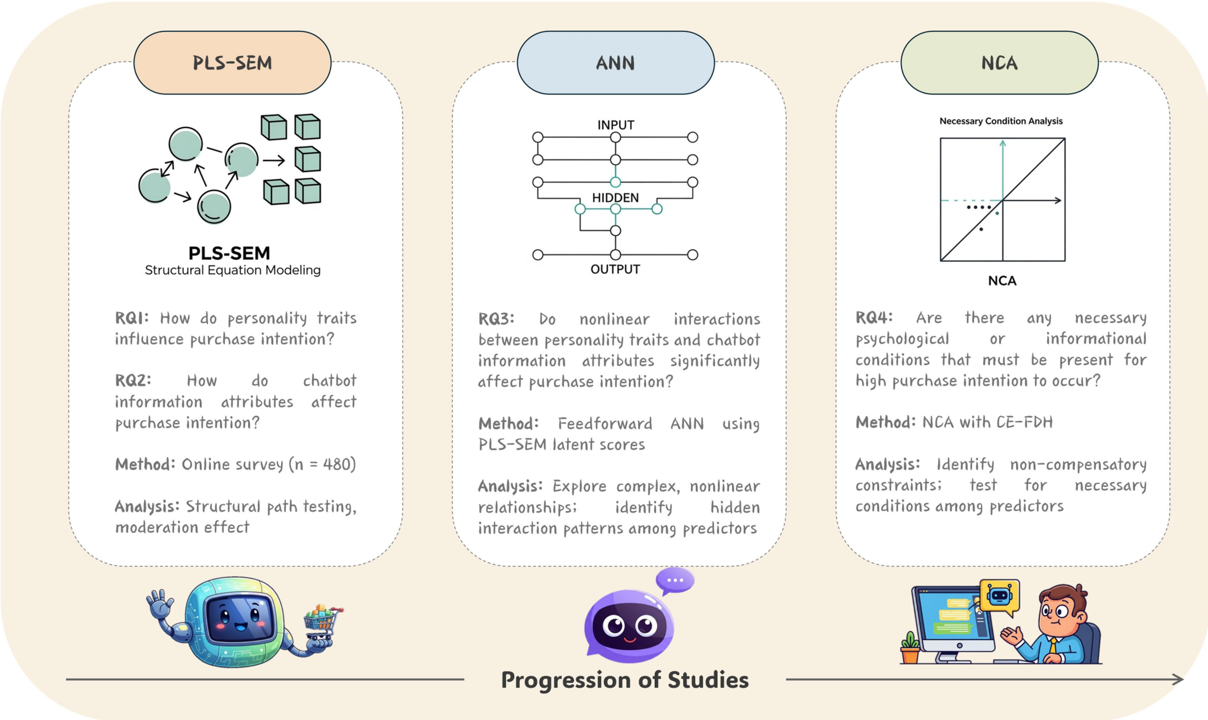


FIGURE 1
Analytical framework combining PLS-SEM, ANN, and NCA.

product-recommendation chatbots have become a crucial tool for enhancing the consumer experience (Rahevar and Darji, 2024). These AI-driven agents not only offer personalized product suggestions but also facilitate highly interactive, human-like dialogs, thereby transforming conventional online shopping into a more engaging and socially enriched experience (Chen et al., 2025). Among diverse consumer segments, Generation Z—digital natives born between 1997 and 2012—demonstrates a particularly high level of openness to AI technologies and a pronounced reliance on algorithmic support in decision-making processes (Bunea et al., 2024). To understand how this cohort responds to chatbot-based product recommendations, it is essential to construct a theoretical framework that encompasses both individual psychological dispositions and users' perceptions of chatbot characteristics.

To this end, the present study draws upon two classical theoretical perspectives—namely, the Big Five personality trait model and source credibility theory—to explain the formation of purchase intention in AI-mediated retail settings. The Big Five model categorizes personality across five dimensions: extraversion, agreeableness, conscientiousness, neuroticism, and openness to experience. These traits have demonstrated robust predictive validity across a variety of digital behaviors, including technology adoption, trust in automated systems, and responsiveness to recommendation algorithms (Seyfi et al., 2025). Prior studies have indicated that although Generation Z is generally receptive to technological innovation, their personality profiles differ significantly, potentially leading to varied patterns of interaction with AI agents (Ding et al., 2025). For instance, individuals with high openness are more likely to embrace novel chatbot interfaces, while those with high conscientiousness tend to pay closer attention to the quality of the information provided (Kovbasiuk et al., 2025).

At the same time, source credibility theory posits that perceived expertise, trustworthiness, and interactivity are key factors influencing how users evaluate information sources (Sardar et al., 2024). In the context of AI chatbots, these attributes are not only embedded in system design but also subjectively interpreted by users during the interaction process (Anbalagan et al., 2025). When a chatbot is perceived as competent, reliable, and engaging, its recommendations are more likely to be viewed as persuasive and actionable (Kim and Priluck, 2025). For Generation Z, characterized by high digital literacy and extensive exposure to mediated content, these source-level cues may be especially influential in shaping behavioral outcomes (Ding et al., 2025).

The integration of these two theoretical perspectives contributes to a more comprehensive understanding of how purchase intentions are formed in chatbot-mediated contexts. The Big Five framework elucidates stable intrapersonal differences in psychological predisposition, while source credibility theory reveals how users cognitively and affectively respond to external system cues during interaction. This integrative framework allows for the simultaneous modeling of internal psychological drivers and external perceptual influences on consumer decision-making, which is particularly well-suited for uncovering the dual-layered mechanisms by which Generation Z users engage with AI-based recommendation agents. As such, the integration of these theories aligns with the complex and interactive nature of technology-mediated consumption and directly addresses this study's dual emphasis on personality-driven variability and source-driven persuasion.

3.1 Big five personality traits

Individual cognition and behavior may be shaped by various factors, including personal experiences, living environment, and educational attainment (Baumert et al., 2017; Salmony and Kanbach, 2022). Personality is widely regarded as a foundational psychological construct that influences how individuals perceive, feel, and act (Sadeq et al., 2018; Louwen et al., 2023). It is jointly determined by genetic predispositions and socio-cultural factors. While many personality traits have a hereditary basis, social environments contribute significantly to their formation, guiding individuals toward similar personality structures (Hörz-Sagstetter et al., 2021). Personality encapsulates a distinct configuration of thoughts, emotions, and behaviors that distinguishes one individual from another, reflecting consistent and coherent psychological attributes (Boudreaux, 2016). Individual personality traits have been shown to predict and explain specific behavioral tendencies (Zweig and Webster, 2003; Busseri and Erb, 2024), including technology adoption behaviors (Liu et al., 2024). Examining new technology usage from the perspective of personality traits offers deeper insights into the psychological mechanisms underlying individual acceptance. Prior research indicates that personality traits significantly affect individuals' adoption of internet platforms (Svendsen et al., 2013), social media (Liu and Campbell, 2017), metaverse applications (Kumar et al., 2024), and artificial intelligence technologies (Riedl, 2022).

Personality is a multifaceted system encompassing temperament, character, cognitive styles, and self-regulatory processes. It is marked by distinctiveness, consistency, comprehensiveness, and adaptability (Haslam et al., 2004; Cuartero and Tur, 2021). Trait theory posits that personality consists of a constellation of traits—stable, cross-situational patterns of behavior and thought (Bleidorn et al., 2016). Traits are defined as enduring dispositions with moderate temporal stability, capturing individual preferences and behavioral tendencies (Fajkowska, 2017; Hopwood, 2025). The big five personality traits—extraversion, agreeableness, conscientiousness, neuroticism, and openness—are the most widely accepted framework in contemporary psychology (Lynn, 2021).

Extraversion describes individuals' social energy and activity levels. It embodies enthusiasm, sociability, and assertiveness, reflecting the extent of engagement in interpersonal interactions (Buecker et al., 2020). Extraverts typically thrive in group settings and enjoy engaging in new experiences. In contrast, introverts prefer solitude and derive energy from introspection (Eid et al., 2003). Due to their curiosity and exploratory tendencies, extraverts are more likely to engage with chatbots and respond positively to product recommendations.

Agreeableness captures traits related to altruism, empathy, and cooperation. Individuals high in agreeableness are typically considerate, trusting, and inclined toward social harmony (Wilmot and Ones, 2022). In contrast, those low in agreeableness may exhibit greater skepticism and competitiveness (Stavrova et al., 2022). Agreeable individuals tend to respond positively to personalized experiences; hence, chatbot recommendations that align with their interpersonal orientation may increase purchase intentions.

Conscientiousness reflects goal-directed behavior and self-discipline. It is associated with organization, responsibility, and persistence (Buelow and Cayton, 2020). High-conscientiousness individuals are more likely to engage in planned, deliberate decision-making. They may place greater trust in reliable and well-structured chatbot systems, perceiving them as extensions of efficient service. Conversely, low-conscientiousness individuals may prefer spontaneity and demonstrate less concern for rule-based interactions (Fleischmann et al., 2023).

Neuroticism refers to emotional instability and sensitivity to stress (Asaoka et al., 2020). Individuals scoring high in neuroticism are more prone to negative emotions, such as anxiety and mood swings, while emotionally stable individuals tend to be calm and resilient (Lee and Bottomley, 2021). In a consumer context, neuroticism may influence the emotional reactions to chatbot interactions and increase the likelihood of impulse purchases under stress or emotional arousal.

Openness denotes intellectual curiosity, creativity, and a preference for novelty (Tucaković and Nedeljković, 2022). Open individuals are more receptive to new experiences and are more likely to explore unfamiliar options (Gil de Zúñiga et al., 2017). Consequently, they may be more engaged with innovative chatbot functionalities and willing to accept AI-generated product recommendations. These individuals are also more adept at integrating new information into their decision-making processes.

Therefore, the following hypothesis is proposed:

> *H1:* Extraversion has a significant positive impact on Generation Z's intention to purchase products recommended by product-recommendation chatbots.
>
> *H2:* Agreeableness has a significant positive impact on Generation Z's intention to purchase products recommended by product-recommendation chatbots.
>
> *H3:* Conscientiousness has a significant positive impact on Generation Z's intention to purchase products recommended by product-recommendation chatbots.
>
> *H4:* Neuroticism has a significant positive impact on Generation Z's intention to purchase products recommended by product-recommendation chatbots.
>
> *H5:* Openness has a significant positive impact on Generation Z's intention to purchase products recommended by product-recommendation chatbots.

3.2 Information source characteristics

In societal contexts, information encompasses a wide array of content that serves as a ubiquitous medium, linking individuals with knowledge about people, events, and objects. Within the sphere of electronic commerce, information plays a critical role (Zhong and Han, 2023). As the vehicle for disseminating content, the information source determines both the quality and substance of that content, rendering its role indispensable. The trustworthiness and effectiveness of information sources have been extensively examined in theoretical models, which explain how various source attributes influence audience attitudes and behaviors (Kelman, 2017). Through processes such as production, processing, storage, and dissemination, information reaches its recipients and exerts persuasive effects—an influence regulated largely by the characteristics of the information source. These persuasive effects shape recipients' attitudes, cognitive perceptions, and behavioral outcomes, most notably influencing advertising attitudes, brand perceptions, and purchase intentions (Huete-Alcocer et al., 2019).

Expertise, as a key attribute of an information source, reflects the degree to which audiences perceive it as capable of delivering accurate and relevant expertise. A highly professional information source provides audiences with reliable, in-depth knowledge, thereby fostering psychological compliance and trust (Dong et al., 2023). This perceived expertise can shape consumer attitudes positively (Liu et al., 2023). Product recommendation chatbots, for instance, rely on advanced algorithms and vast user data to analyze consumer behavior and generate tailored product suggestions. These features contribute to the perceived precision and expertise of chatbot recommendations. Moreover, many chatbots possess learning capabilities that allow them to continuously refine their algorithms based on user feedback and behavioral data, enhancing both the accuracy and trustworthiness of their suggestions (Haugeland et al., 2022). Consequently, consumers may perceive such chatbots as both professional and trustworthy, fostering more favorable attitudes toward the recommended products and increasing their purchase intentions.

Trustworthiness, another critical attribute of an information source, pertains to the perceived trustworthiness of the content provider. When facing important decisions, individuals typically seek out credible sources to obtain relevant and reliable information, which helps reduce perceived risks and uncertainty (Erdogan, 1999). Chatbots that analyze users' historical behaviors and preferences can offer highly personalized recommendations, increasing users' trust. In addition, the consistency and data-driven nature of chatbot recommendations—free from many human biases—further contribute to their perceived trustworthiness (Lin et al., 2020). As such, credible chatbots reduce consumers' informational uncertainty, filling knowledge gaps and acting as reliable reference points during the decision-making process.

Interactivity refers to the extent and quality of real-time communication between the information source and the audience, thereby creating a sense of social presence (He et al., 2022). Chatbots commonly leverage advanced natural language processing (NLP) technologies to interpret and engage with natural language inputs from users (Attigeri et al., 2024). This enables them to conduct fluid and dynamic conversations that enhance perceived interactivity. Additionally, chatbots often deliver real-time feedback, which further facilitates interactive dialog (Tsai et al., 2021). This real-time responsiveness allows users to ask questions, share opinions, or request additional information, deepening their engagement. By fostering two-way communication, chatbot interactivity not only enhances users' understanding of the products but also increases their involvement in the decision-making process (Meier et al., 2024). The more engaged users are, the more likely they are to develop favorable attitudes and make informed purchase decisions.

Customization stands as a cornerstone of product recommendation chatbot functionality within the e-commerce domain. This attribute refers to a chatbot's ability to tailor its responses and product suggestions to fit the specific preferences and needs of individual users (Skjuve et al., 2021). Such adaptive capacity enables chatbots to serve diverse consumer groups effectively across various application scenarios. Through AI-driven analytics, chatbots swiftly interpret user data to predict interest areas and deliver personalized recommendations (Jiang et al., 2023). This tailored engagement not only enhances operational efficiency but also nurtures stronger consumer–brand relationships. Advanced customization further enables chatbots to monitor and learn from user behavior in real time, refining their ability to anticipate needs with greater accuracy (Wang C. et al., 2023; Wang L. et al., 2023; Wang X. et al., 2023). By dynamically adjusting their outputs to align with individual user profiles, chatbots elevate the quality of user interaction and overall satisfaction. Thus, AI-enabled customization enriches the consumer experience and strengthens the persuasive power of chatbot interactions, ultimately influencing user preferences and purchase intentions.

Therefore, the following hypothesis is proposed:

> *H6:* Expertise has a significant positive impact on the intention of Generation Z to purchase products recommended by product recommendation chatbots.
>
> *H7:* Interactivity has a significant positive impact on the intention of Generation Z to purchase products recommended by product recommendation chatbots.
>
> *H8:* Trustworthiness has a significant positive impact on the intention of Generation Z to purchase products recommended by product recommendation chatbots.
>
> *H9:* Customization has a significant positive impact on the intention of Generation Z to purchase products recommended by product recommendation chatbots.

3.3 Personal innovativeness

Personal innovativeness refers to an individual's propensity to embrace novel ideas, methods, and technologies (Walley et al., 2017). It reflects one's willingness to take risks and adapt to change. Individuals exhibiting high levels of innovativeness are typically driven to seek out experiences that are new and unconventional (Ali and Warraich, 2023). Such individuals are more inclined to explore emerging products and technologies as a means of satisfying their curiosity and pursuit of novelty (Alkawsi et al., 2021). As chatbot-mediated shopping represents a novel digital experience, it is likely to appeal to highly innovative consumers who demonstrate greater openness to technological experimentation. Furthermore, these individuals often possess a heightened capacity to adapt to and adopt new technologies (Venkatesh et al., 2012). They may find it easier to comprehend and engage with the services offered by chatbots and exhibit greater willingness to purchase chatbot-recommended products. Their receptivity to innovation enhances their ability to recognize and appreciate the convenience and benefits enabled by advanced technologies.

Accordingly, the following hypothesis is proposed:

> *H10:* Personal innovativeness positively moderates the influence of personality traits on the intention of Generation Z to purchase products recommended by product recommendation chatbots.

Consumer decision-making in chatbot-assisted environments is a result of the interaction between internal psychological predispositions and external information stimuli. The Big Five personality traits serve as a theoretically grounded framework for capturing stable interindividual differences in cognition, emotion, and behavioral tendencies. These traits influence how consumers attend to, interpret, and react to external cues in digital environments. For example, traits such as openness or neuroticism may alter the way individuals perceive risk, trustworthiness, or novelty when interacting with artificial agents. On the other hand, information source characteristics—such as perceived expertise, trustworthiness, and interactivity—represent functional cues emitted by the chatbot that guide users' evaluations of message quality and decision relevance. The decision to integrate these two dimensions in a single model stems from the recognition that behavioral outcomes such as purchase intention are not merely functions of either user traits or system attributes alone, but emerge from their interplay.

This theoretical integration enables the investigation of how personality-based perceptual filters modulate responses to chatbot-generated recommendations. It assumes that the influence of information cues is not uniform across individuals but is differentially processed depending on who the user is. In this sense, personality traits act as endogenous filters that shape the salience and interpretive meaning of chatbot characteristics. For instance, while interactivity may enhance purchase intention among extraverted users who enjoy social-like interaction, the same feature may be less effective or even distracting for individuals high in conscientiousness who prioritize efficiency and clarity.

In addition, the inclusion of personal innovativeness as a moderating variable further refines the model by accounting for variability in users' openness to novel technologies. Even among individuals with similar personality profiles, those with higher levels of innovativeness are more likely to engage with and respond positively to chatbot-driven interaction, especially when the chatbot exhibits high trustworthiness or advanced interactive capabilities. By modeling personal innovativeness as a moderator, the framework captures the boundary conditions under which internal traits and external cues jointly translate into behavioral intentions. Therefore, the combination of personality traits, information source characteristics, and innovativeness orientation provides a comprehensive structure for understanding individual-level variability in AI-mediated consumer behavior. The research hypothesis model is shown in Figure 2.

4 Research methodology

4.1 Measurement

This study employed a structured survey questionnaire to empirically evaluate the proposed research model. To ensure the validity and reliability of the measurement, each construct in the model was operationalized using multiple items adapted from well-established scales in prior literature. These items were carefully modified to align with the specific research context. As outlined in Table 1, the measurement framework included items for purchase intention, drawn from Pavlou (2003), and dimensions of product recommendation chatbots, informed by Ohanian (1991), Gorham (1988), Yang et al. (2023), and Periaiya and Nandukrishna (2023). The chatbot-related constructs were conceptualized across four dimensions: expertise, trustworthiness, interactivity, and customization. The Big Five personality traits were measured using items based on the scale developed by Benet-Martínez and John (1998), while personal innovativeness was assessed using items primarily adapted from Agarwal and Prasad (1999). To ensure consistent interpretation of trait direction, the measurement items for Neuroticism were reverse-coded prior to analysis, as indicated in Table 1. All measurement items employed a seven-point Likert scale, ranging from 1 ("strongly disagree") to 7 ("strongly agree"), enabling

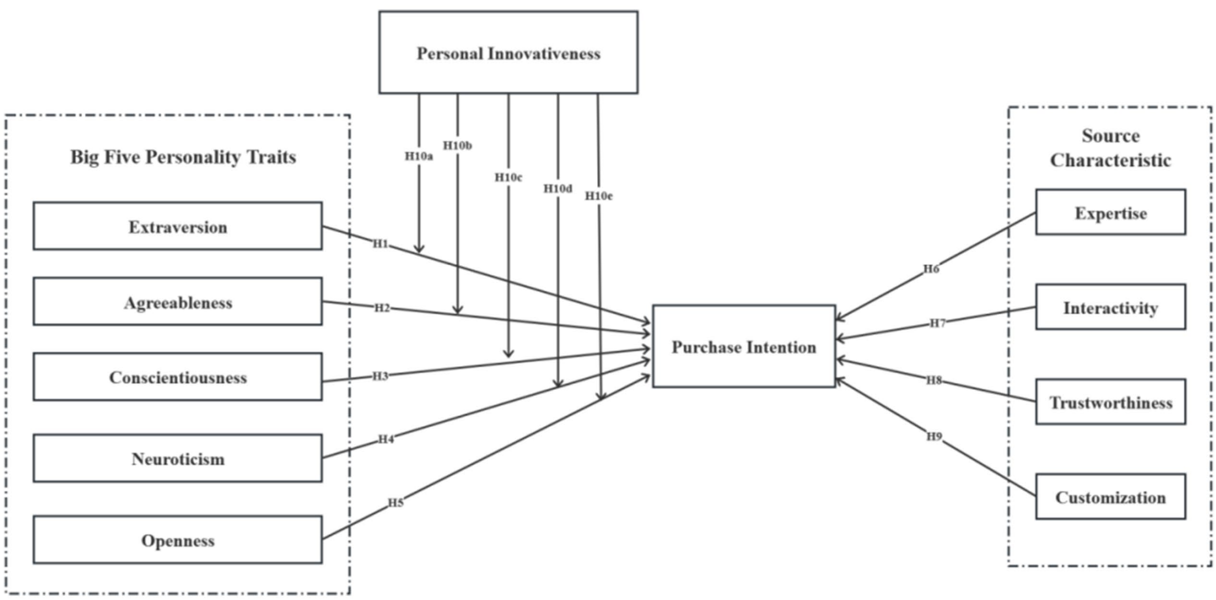


FIGURE 2
Research model.

TABLE 1 Measurement scales.

Variable	Items	Source
Expertise	I believe that the product recommendation chatbot I use possesses professional skills	Yang et al. (2023), Ohanian (1991), and Gorham (1988)
	I think the product recommendation chatbot I use has special skills and expertise	
	I believe the product recommendation chatbot I use is knowledgeable	
	I feel that the product recommendation chatbot I use has extensive experience in recommending products	
Trustworthiness	I believe the products recommended by the product recommendation chatbot I use are trustworthy	
	I trust the product recommendation chatbot I use	
	I believe the content recommended by the product recommendation chatbot I use is reliable	
Interactivity	I believe I have a good interactive relationship with the product recommendation chatbot I use	
	I feel that the content recommended by the product recommendation chatbot I use allows me to engage effectively	
	I think the content recommended by the product recommendation chatbot I use can pique my interest	
Customization	I believe the product recommendation chatbot shows me customized content	Periaiya and Nandukrishna (2023)
	I feel the product recommendation chatbot is tailored for my use	
	I think the products recommended by the product recommendation chatbot are tailored for my use	
Purchase intention	If given the opportunity, I plan to purchase the products recommended by the product recommendation chatbot.	Pavlou (2003)
	If given the opportunity, I predict that I will purchase the products recommended by the product recommendation chatbot in the future.	
	I am highly likely to purchase the products recommended by the product recommendation chatbot in the near future.	

(Continued)

10.3389/fpsyg.2025.1454197

TABLE 1 (Continued)

Variable	Items	Source
Extraversion	I feel I am outgoing and sociable person	Benet-Martínez and John (1998)
	I am very talkative	
	I have an assertive personality	
	I usually generate a lot of enthusiasm	
Agreeableness	I am considerate to almost everyone	
	I like to cooperate with others	
	I am always helpful and unselfish with others	
	I have a forgiving nature	
Conscientiousness	I will do job thoroughly	
	I do things efficiently	
	I stick to my plans	
	I am a reliable person	
Neuroticism	I do not worry a lot	
	I never get tensed	
	I do not get nervous easily	
	I generally remain calm in tense situations	
Openness	I am more inventive	
	I am open to new ideas	
	I feel I have active imagination	
	I like to reflect and play with ideas	
	I love art, music and literature	
	I am a deep thinker	
	I am curious about many different things	
	I prefer to do works that is challenging	
Personal innovativeness	I believe I am ready and capable of using innovative technologies such as product purchase in product recommendation chatbots	Agarwal and Prasad (1999)
	When I hear about new information technology I would look for ways to experiment with it	
	I like to experiment with new IT products	
	Among my peers, I am usually the first to try IT products	

10.3389/fpsyg.2025.1454197

respondents to express the degree of their agreement with each statement.

A seven-point Likert scale was selected because prior psychometric research shows that scales with 5–7 response categories maximize reliability, item discrimination, and respondent preference while avoiding the cognitive overload associated with longer formats (Preston and Colman, 2000; Finstad, 2010). In the context of consumer-behavior surveys, a seven-point format provides finer granularity than a five-point scale yet retains cross-cultural comparability (Dawes, 2008). Recent chatbot and technology-adoption studies have likewise adopted seven-point scales for the same reasons of sensitivity and ease of interpretation (Yang et al., 2023; Periaiya and Nandukrishna, 2023). Therefore, using a seven-point Likert scale aligns with best practice and enhances the psychometric quality of our measurements.

In addition, the questionnaire included demographic variables such as age, gender, occupation, and place of residence. Following the initial compilation of the survey instrument, the questionnaire was translated into Chinese by a native Chinese-speaking researcher to ensure linguistic and contextual appropriateness. To further enhance clarity and accuracy, five graduate students in management—each with prior experience using product recommendation chatbots—were invited to review and provide feedback on the translated version. Subsequently, the revised Chinese questionnaire was back-translated into English by a researcher with expertise in academic English to ensure semantic equivalence and consistency across both language versions.

4.2 Data collection and descriptive statistics

This study conducted an online questionnaire survey targeting Generation Z consumers in China (aged 16–29) to empirically examine the proposed research hypotheses. Several factors justify the focus on the Chinese context. First, China boasts an enormous online population of approximately 1.092 billion, with an internet penetration rate of 77.5%. A significant portion of these users belong to Generation Z. Second, product recommendation chatbots are widely implemented in Chinese e-commerce platforms and have become an integral part of online shopping experiences.

The questionnaire was designed and administered using Wenjuanxing,[1] the largest and most widely utilized online survey platform in China. Wenjuanxing is frequently adopted by both domestic and international enterprises as well as academic institutions due to its robust sampling capabilities and efficient data collection infrastructure. To obtain a sufficient number of eligible respondents, a convenience sampling approach was employed. The survey link was disseminated through three major Chinese social media platforms—Sina Weibo, Tencent WeChat, and Douyin (TikTok China)—via posts, private messages, and group announcements. As an incentive, participants who completed the survey received a nominal monetary reward of 5 RMB. Prior to the main questionnaire, two mandatory screening questions were used to ensure respondent eligibility: (a) "Are you currently aged 16–29 (Generation Z)?" and (b) "Have you ever used product recommendation services provided by chatbots?" Respondents who answered "No" to either question were automatically excluded from the survey. Eligible participants then proceeded to complete the full set of measurement items covering all constructs of interest: Big Five personality traits, perceived information source characteristics (expertise, trustworthiness, interactivity, and customization), personal innovativeness, and purchase intention, along with demographic information.

A total of 683 responses were collected. After excluding questionnaires from individuals who did not meet the age criterion, had not used chatbot-based product recommendations, provided incomplete responses, completed the survey in less than 1 min, or selected the same option for all items, a final valid sample of 480 responses was retained for analysis. Descriptive statistics for the sample are summarized in Table 2. Regarding age, the respondents ranged from 16 to 29 years old, with a mean of 21.86 years and a standard deviation of 2.41. In terms of gender, 73.33% ($n = 352$) identified as male and 26.67% ($n = 128$) as female. Regarding occupation, 72.71% ($n = 349$) were students and 27.29% ($n = 131$) were employed professionals. With respect to residence, 82.29% ($n = 395$) reported living in urban areas, while 17.71% ($n = 85$) resided in rural areas.

5 Results

5.1 Common method bias

Common method bias (CMB) refers to systematic, non-substantive variance that arises from measurement artifacts—such as questionnaire format, respondent characteristics, or contextual influences—rather than the constructs of interest themselves (Chang and Zhu, 2012). It represents a critical threat to the validity of empirical findings derived from self-reported survey data. To assess the presence of CMB in this study, Harman's single-factor test was conducted following the procedure recommended by Podsakoff et al. (2003). An exploratory factor analysis was performed using SPSS 27.0. The results revealed that 11 factors had eigenvalues greater than 1,

TABLE 2 Demographic information.

Demographic measures	Count	Percentage
Gender		
Female	128	26.67%
Male	352	73.33%
Occupation		
Students	349	72.71%
Working professionals	131	27.29%
Place of living		
Urban	395	82.29%
Rural	85	17.71%

1 https://www.wjx.cn/

collectively accounting for 71.90% of the total variance. Notably, the first unrotated factor accounted for only 27.78% of the variance, which falls well below the commonly accepted threshold of 40%. These findings suggest that common method bias is not a significant concern in this study.

5.2 Results of structural equation modeling

5.2.1 Assessment of measurement model

This study assessed the reliability of the measurement scale primarily through internal consistency analyses. Specifically, composite reliability (*CR*) and Cronbach's α coefficients were used as indicators. Both *CR* and Cronbach's α values exceeding the threshold of 0.70 are considered indicative of strong reliability. As shown in Table 3, all latent constructs demonstrated *CR* and α values above this benchmark, confirming the reliability of the measurement model.

Validity was evaluated in terms of both convergent and discriminant validity. Convergent validity was assessed through factor loadings, with values above 0.70 deemed acceptable, indicating that the observed variables adequately represent the underlying latent constructs. Discriminant validity was examined using the Fornell–Larcker criterion and the Heterotrait–Monotrait ratio (*HTMT*). According to the Fornell–Larcker criterion, discriminant validity is established when the average variance extracted (*AVE*) for each construct exceeds 0.50 and is greater than the squared correlations between constructs. *HTMT* values below 0.85 (or, in some cases, 0.90) further confirm satisfactory discriminant validity by indicating that correlations within constructs are stronger than those between different constructs.

As presented in Tables 4, 5, all factor loadings met the criteria for structural validity (Fornell and Larcker, 1981), demonstrating a strong linear relationship between observed indicators and their respective latent variables, as well as sufficient explanatory power. Furthermore, to control for potential collinearity issues among predictors, variance inflation factors (*VIF*) were examined. All *VIF* values were below the critical threshold of five, indicating that multicollinearity was not a concern in this model. The multicollinearity diagnostics reported in Tables 6, 7 further support the robustness of the structural model (Neter et al., 1990).

5.2.2 Assessment of structural model

The predictive validity of the research model primarily hinges on the explanatory power of the endogenous variables. The coefficient of determination (R^2) and effect size (f^2) serve as the principal indicators of predictive strength, with f^2 values of 0.020, 0.150, and 0.350 denoting small, medium, and large effect sizes, respectively. The results of the structural model are presented in Table 8 and Supplementary Figure 1.

Extraversion demonstrates a significant positive effect on purchase intention ($\beta = 0.216$, $p < 0.001$), with an effect size of 0.094, indicating a small effect. Agreeableness also shows a significant positive relationship with purchase intention ($\beta = 0.175$, $p < 0.001$), but with a smaller effect size of 0.044. Neither conscientiousness ($\beta = 0.020$, $p > 0.05$) nor neuroticism ($\beta = 0.021$, $p > 0.05$) exert a significant impact on purchase intention. Openness is positively associated with purchase intention ($\beta = 0.149$, $p < 0.01$), with an effect size of 0.022, also reflecting a small effect.

Among the information source characteristic, expertise ($\beta = 0.123$, $p < 0.05$), interactivity ($\beta = 0.097$, $p < 0.01$), and customization ($\beta = 0.194$, $p < 0.01$) all have significant positive effects on purchase intention, with corresponding effect sizes of 0.020, 0.022, and 0.050, respectively—each indicating small effects. Trustworthiness does not exhibit a statistically significant impact ($\beta = 0.056$, $p > 0.05$).

Personal innovativeness shows a significant positive influence on purchase intention ($\beta = 0.329$, $p < 0.001$), with an effect size of 0.223, suggesting a moderate effect. Additionally, the interaction between personal innovativeness and extraversion is significant ($\beta = 0.104$, $p < 0.05$), with an effect size of 0.031. To further examine this moderation effect, Process Macros (Model 1) was employed. As illustrated in Supplementary Figure 2, extraversion has no significant impact on purchase intention at low levels of personal innovativeness ($\beta = 0.120$, $p > 0.05$). At moderate levels of innovativeness, extraversion shows a significant positive effect ($\beta = 0.234$, $p < 0.001$), which intensifies at high levels ($\beta = 0.358$, $p < 0.001$). These findings suggest that the positive influence of extraversion on purchase intention becomes stronger as personal innovativeness increases. In contrast, the interaction terms between personal innovativeness and agreeableness ($\beta = 0.074$, $p > 0.05$), conscientiousness ($\beta = 0.021$, $p > 0.05$), neuroticism ($\beta = 0.074$, $p > 0.05$), and openness ($\beta = 0.018$, $p > 0.05$) are not statistically significant.

Overall, the model accounts for 58.80% of the variance in Generation Z's purchase intention regarding chatbot-recommended products. Furthermore, the blindfolding procedure was applied to evaluate the Q^2-values of the five endogenous constructs. Since all Q^2-values exceeded zero, the model demonstrates satisfactory predictive relevance. Specifically, the Q^2-value for purchase intention ($Q^2 = 0.428$) confirms the model's reliable predictive capability for the endogenous outcome variables.

5.3 ANN results

In recent years, the reliability of structural equation modeling (SEM) has been increasingly questioned due to its inherent assumption of linear and compensatory relationships between constructs. This assumption may oversimplify the complex, multifactorial nature of decision-making processes. Given the exploratory nature of our research domain—which is still emerging and underexplored—a more robust and complementary analytical approach is necessary to validate and extend SEM findings. Accordingly, this study developed an artificial neural network (ANN) model based on the backpropagation (BP) algorithm, where extraversion (EXT), agreeableness (AGR), openness (OPE), expertise (EXP), interactivity (INT), customization (CUS), and personal innovativeness (PEI) were used as input variables, and purchase intention (PI) as the output variable.

The BP neural network minimizes the squared error between predicted and actual values using gradient descent to iteratively update weights and thresholds, optimizing the alignment between predicted outputs and expected outcomes (Denoeux and Lengellé, 1993). Table 9 presents the predictive performance of the ANN model.

To examine potential nonlinear relationships within the model, we compared the predictive performance of PLS-SEM and ANN. As

TABLE 3 Cronbach's α, corporate reliability and average variance extracted.

Variable	Cronbach's alpha	Composite reliability	AVE	Items	Factor loading
Expertise	0.811	0.875	0.636	EXP1	0.772
				EXP2	0.826
				EXP3	0.841
				EXP4	0.748
Personal innovativeness	0.936	0.954	0.838	PEI1	0.914
				PEI2	0.915
				PEI3	0.909
				PEI4	0.923
Interactivity	0.822	0.889	0.729	INT1	0.770
				INT2	0.888
				INT3	0.898
Trustworthiness	0.860	0.915	0.781	TRU1	0.881
				TRU2	0.878
				TRU3	0.892
Agreeableness	0.878	0.916	0.732	AGR1	0.868
				AGR2	0.868
				AGR3	0.835
				AGR4	0.85
Customization	0.855	0.912	0.776	CUS1	0.878
				CUS2	0.903
				CUS3	0.861
Extraversion	0.86	0.905	0.703	EXT1	0.882
				EXT2	0.837
				EXT3	0.805
				EXT4	0.829
Openness	0.876	0.897	0.529	OPE1	0.877
				OPE2	0.825
				OPE3	0.818
				OPE4	0.835
				OPE5	0.764
				OPE6	0.764
				OPE7	0.876
				OPE8	0.864
Neuroticism	0.829	0.885	0.659	NEU1	0.820
				NEU2	0.850
				NEU3	0.828
				NEU4	0.745
Conscientiousness	0.946	0.961	0.860	CON1	0.930
				CON2	0.926
				CON3	0.942
				CON4	0.910
Purchase intention	0.853	0.911	0.773	PI1	0.901
				PI2	0.880
				PI3	0.856

EXP, Expertise; PEI, Personal Innovativeness; INT, Interactivity; TRU, Trustworthiness; EXT, Extraversion; CUS, Customization; AGR, Agreeableness; OPE, Openness; NEU, Neuroticism; CON, Conscientiousness; PI, Purchase Intention.

TABLE 4 Differential validity based on HTMT method.

Variable	EXP	PEI	INT	TRU	AGR	CUS	EXT	OPE	NEU	CON	PI
EXP											
PEI	0.186										
INT	0.106	0.036									
TRU	0.636	0.169	0.034								
AGR	0.538	0.241	0.111	0.490							
CUS	0.645	0.303	0.073	0.553	0.563						
EXT	0.210	0.140	0.069	0.309	0.295	0.379					
OPE	0.696	0.216	0.094	0.613	0.590	0.601	0.332				
NEU	0.550	0.224	0.102	0.512	0.476	0.543	0.326	0.574			
CON	0.047	0.037	0.099	0.029	0.018	0.044	0.126	0.108	0.056		
PI	0.558	0.529	0.173	0.429	0.605	0.654	0.470	0.570	0.457	0.058	

EXP, Expertise; PEI, Personal Innovativeness; INT, Interactivity; TRU, Trustworthiness; EXT, Extraversion; CUS, Customization; AGR, Agreeableness; OPE, Openness; NEU, Neuroticism; CON, Conscientiousness; PI, Purchase Intention.

TABLE 5 Fornell-Larcker criterion.

Variable	EXP	PEI	INT	TRU	AGR	CUS	EXT	OPE	NEU	CON	PI
EXP	0.798										
PEI	0.171	0.915									
INT	0.087	0.023	0.854								
TRU	−0.526	−0.151	−0.024	0.884							
AGR	0.459	0.219	0.101	−0.425	0.855						
CUS	0.534	0.273	0.065	−0.475	0.488	0.881					
EXT	0.185	0.127	0.025	−0.268	0.259	0.328	0.839				
OPE	0.631	0.226	0.087	−0.585	0.575	0.558	0.308	0.727			
NEU	−0.453	−0.204	−0.086	0.437	−0.413	−0.458	−0.282	−0.547	0.812		
CON	0.042	0.026	−0.094	−0.017	−0.003	0.023	0.114	0.076	−0.042	0.927	
PI	0.477	0.474	0.156	−0.369	0.525	0.56	0.408	0.544	−0.393	0.054	0.879

EXP, Expertise; PEI, Personal Innovativeness; INT, Interactivity; TRU, Trustworthiness; EXT, Extraversion; CUS, Customization; AGR, Agreeableness; OPE, Openness; NEU, Neuroticism; CON, Conscientiousness; PI, Purchase Intention.

recommended by Lau et al. (2021), if the ANN demonstrates superior goodness of fit relative to the linear model, this implies the presence of nonlinear patterns. The average prediction accuracy of the ANN across training and test sets ranged from 73.415 to 85.127%, with an overall average of 80.864%. In contrast, the linear PLS-SEM model accounted for 58.80% of the variance in the outcome variable (see Supplementary Figure 2). These results suggest that the nonlinear BP neural network provides a better fit to the data, thereby confirming the existence of nonlinear relationships among variables in the conceptual model.

Subsequently, a sensitivity analysis was conducted using the permutation method to rank the relative importance of each input variable. As shown in Table 10 and Supplementary Figure 3, customization (CUS, 0.217) emerged as the most influential predictor of purchase intention in the ANN model, followed closely by personal innovativeness (PEI, 0.213). These findings underscore the central role of customization in shaping Generation Z consumers' purchasing behavior in chatbot interactions. Generative AI technologies underpinning advanced customization capabilities can accurately infer and adapt to user preferences, thereby optimizing user experiences. Through tailored recommendations, content delivery, and real-time feedback, AI systems elevate user engagement and facilitate seamless decision-making, ultimately increasing purchase likelihood among digitally savvy Generation Z users. Notably, individuals with high personal innovativeness are more open to adopting new technologies, including product recommendation chatbots, and are thus more likely to act on the product suggestions provided.

According to Table 11, in the PLS-SEM analysis, personal innovativeness was identified as the most influential psychological trait affecting purchase intention, while customization was ranked third among the chatbot features. In contrast, the ANN analysis identified customization as the most critical predictor, with personal innovativeness ranking second. This divergence arises from the different assumptions underlying the two models. The linear PLS-SEM assumes additive and independent effects of predictors, potentially underestimating the interactive or synergistic effects of features such as customization. In such models, personal innovativeness may appear to dominate as a stable, direct influence on behavior, while customization's role is diluted.

TABLE 6 Collinearity test results for outer model (VIF).

Items	VIF	Items	VIF	Items	VIF	Items	VIF	Items	VIF
EXP1	1.543	INT3	2.038	CUS3	1.955	OPE6	2.321	CON4	3.979
EXP2	1.770	TRU1	2.169	EXT1	2.282	OPE7	1.879	PI1	2.363
EXP3	1.749	TRU2	2.175	EXT2	2.084	OPE8	1.474	PI2	2.114
EXP4	1.614	TRU3	2.182	EXT3	1.836	NEU1	1.800	PI3	1.949
PEI1	3.578	AGR1	2.396	EXT4	1.933	NEU2	1.977	PEI × OPE	1.000
PEI2	3.404	AGR2	2.423	OPE1	3.067	NEU3	1.838	PEI × CON	1.000
PEI3	3.384	AGR3	2.028	OPE2	2.548	NEU4	1.679	PEI × NEU	1.000
PEI4	4.101	AGR4	2.108	OPE3	2.232	CON1	4.112	PEI × EXT	1.000
INT1	1.738	CUS1	2.159	OPE4	2.786	CON2	3.990	PEI × AGR	1.000
INT2	1.846	CUS2	2.401	OPE5	2.492	CON3	4.445		

EXP, Expertise; PEI, Personal Innovativeness; INT, Interactivity; TRU, Trustworthiness; EXT, Extraversion; CUS, Customization; AGR, Agreeableness; OPE, Openness; NEU, Neuroticism; CON, Conscientiousness; PI, Purchase Intention.

TABLE 7 Collinearity test results for inner model (VIF).

Variable	PI
PI	
EXP	1.956
PEI	1.180
INT	1.036
TRU	1.733
AGR	1.681
CUS	1.817
EXT	1.208
OPE	2.510
NEU	1.577
CON	1.045
PI × AGR	1.881
PI × EXT	1.298
PI × CON	1.107
PI × NEU	2.461
PI × OPE	3.126

By contrast, ANN is capable of capturing complex, nonlinear interactions and interdependencies among predictors. As a result, the nonlinear model identifies customization as the most critical feature, reflecting its intricate and dynamic relationship with purchase behavior—particularly when moderated by users' levels of personal innovativeness. Thus, the superior explanatory power of customization in the ANN model highlights its central role in influencing Generation Z's purchasing decisions in chatbot-driven e-commerce environments.

Furthermore, when comparing the linear and nonlinear models, the ranking of most predictors remained consistent, with noticeable changes only observed in customization (CUS), personal innovativeness (PEI), and extraversion (EXT). This consistency reinforces the overall stability and robustness of the model's predictive structure across analytical approaches. The application of ANN offers enhanced precision in capturing nonlinear dynamics and deeper insights into the complex mechanisms driving Generation Z's purchasing behavior in response to chatbot recommendations.

5.4 Results of NCA

After identifying the relative importance of relationships among variables within the research model, the next step involves assessing their necessity. If all relationships are deemed necessary, their *p*-values should be less than 0.05, as suggested by Dul et al. (2020). This implies that each exogenous construct in the model serves as a necessary condition for the occurrence of its respective outcome. Necessary Condition Analysis (NCA) provides a suitable framework for this assessment by evaluating the extent to which a condition must be present for an outcome to occur.

Depending on the nature of the variables, NCA employs two main upper-bound analytical techniques: Ceiling Envelopment (CE) for binary or discrete variables with fewer than five levels, and Ceiling Regression (CR) for discrete or continuous variables with five or more levels. Given that the variables in this study are predominantly continuous or multi-level discrete, CR is more appropriate and is therefore the primary method used for interpretation.

NCA evaluates necessity based on two key indicators: the necessity effect size (d) and the significance level derived from Monte Carlo permutation testing. The effect size d ranges from 0 to 1, with higher values indicating a greater degree of necessity. A statistically significant p-value ($p \leq 0.05$) confirms that the predictor is a necessary condition for the outcome variable. As shown in Supplementary Table 1, Agreeableness, Openness, Expertise, Interactivity, and Customization all exhibit significant necessity effects for purchase intention. In contrast, Extraversion shows a CR effect size of 0.042 with a p-value of 0.376, suggesting it is not a necessary condition ($p > 0.05$). Similarly, Personal Innovativeness has a CR effect size of 0.014 with a p-value of 0.336, indicating it does not qualify as a necessary condition either.

To further explore how varying levels of conditional factors influence purchase intention, a bottleneck analysis was conducted (see Supplementary Table 2). This analysis determines the minimum level each conditional factor must attain to achieve specific thresholds of purchase intention, ranging across the observed spectrum.

TABLE 8 A summary of the PLS path analysis.

PLS path	Path coefficient	T statistics	*p*-value	95% confidence interval	
				Lower bound	Upper bound
Extraversion → purchase intention	0.216	4.515	0.000	0.116	0.306
Agreeableness → purchase intention	0.175	3.504	0.000	0.078	0.275
Conscientiousness → purchase intention	0.020	0.708	0.479	−0.039	0.074
Neuroticism → purchase intention	0.021	0.440	0.660	−0.078	0.116
Openness → purchase intention	0.149	2.612	0.009	0.042	0.265
Expertise → purchase intention	0.123	2.513	0.012	0.028	0.220
Interactivity → purchase intention	0.097	3.011	0.003	0.032	0.161
Trustworthiness → purchase intention	0.056	1.142	0.253	−0.049	0.148
Customization → purchase intention	0.194	2.658	0.008	0.044	0.335
Personal innovativeness → purchase Intention	0.329	6.272	0.000	0.221	0.425
Personal innovativeness × extraversion → purchase intention	0.104	1.966	0.045	0.020	0.223
Personal innovativeness × agreeableness → purchase intention	0.074	1.868	0.062	−0.007	0.152
Personal innovativeness × conscientiousness → purchase intention	−0.021	0.680	0.496	−0.073	0.045
Personal innovativeness × neuroticism → purchase intention	0.074	1.436	0.151	−0.031	0.172
Personal innovativeness × openness → purchase intention	0.018	0.380	0.704	−0.083	0.105

The results indicate that to achieve a 50% purchase intention, the required levels for Agreeableness, Openness, Expertise, and Interactivity are 25.00, 28.57, 25.00, and 23.81%, respectively, while Customization remains nonessential (0%). At the 80% purchase intention level, the required levels rise to 28.57% for Openness, 35.71% for Expertise, 23.81% for Interactivity, and 19.04% for Customization, while Agreeableness remains constant at 25%. At the highest threshold (80%), the required levels for Openness, Expertise, Interactivity, and Customization further increase to 37.50, 50.00, 33.33, and 23.81%, respectively, with Agreeableness still unchanged at 25%.

These findings suggest that at lower thresholds of purchase intention, Agreeableness, Openness, Expertise, and Interactivity function as more critical necessary conditions, whereas the role of Customization is minimal. However, as the desired level of purchase intention increases, the importance of Openness, Expertise, Interactivity, and Customization as necessary conditions becomes more pronounced. Notably, the required level of Agreeableness remains constant across all thresholds, indicating its consistent role as a foundational necessity in driving purchase intention.

6 Discussion

Over the past two decades, the rapid advancement of technology has profoundly transformed daily life. The continuous evolution of artificial intelligence (AI) has brought about substantial changes in education, consumption, social interaction, and culture. Recently, the emergence of chatbots has begun to exert a significant influence on consumer behavior. Generation Z, as the primary users of digital technology, plays a crucial role in shaping consumer behavior through interactions with product recommendation chatbots. Accordingly, this study investigates the key determinants—namely, personality traits

TABLE 9 Accuracy values for neural network.

Neural network	Input: EXT, AGR, OPE, EXP, INT, CUS, PEI Output: PI	
	Training (%)	Testing (%)
ANN1	83.480	81.663
ANN2	80.467	80.364
ANN3	83.610	76.042
ANN4	79.303	73.415
ANN5	81.116	82.692
ANN6	80.795	80.272
ANN7	79.253	85.227
ANN8	78.355	80.167
ANN9	83.183	83.785
ANN10	85.127	78.966
Mean	81.469	80.259
SD	2.257	3.517

EXP, Expertise; PEI, Personal Innovativeness; INT, Interactivity; EXT, Extraversion; CUS, Customization; AGR, Agreeableness; OPE, Openness; PI, Purchase Intention.

and chatbot characteristics—that influence Generation Z consumers' purchase of chatbot-recommended products.

The structural equation modeling (SEM) results reveal that the Big Five personality traits significantly predict the purchase intentions of Generation Z consumers. Extraversion, reflecting social activity, enthusiasm, and openness, was positively associated with engagement. Highly extraverted consumers are typically more responsive to social influence, including peer, influencer, and chatbot-based recommendations (Marengo et al., 2020). Their comfort with new technologies enhances their receptivity to chatbot interactions and increases their likelihood of accepting suggested products (Gan, 2016). Agreeableness, which encompasses cooperativeness, trust, empathy, and friendliness, was also positively associated with purchase intention. Consumers high in agreeableness are more inclined to trust recommendation systems, perceiving chatbot suggestions as supportive and well-intentioned (Sowmya et al., 2023). These consumers tend to appreciate personalized services, and chatbot customization addresses their desire for tailored experiences (Fazli-Salehi et al., 2021). Openness, which captures creativity, curiosity, and receptiveness to novelty, positively influences consumers' willingness to explore and adopt new products and technologies (Tucaković and Nedeljković, 2022; Duong, 2021). Consumers high in openness are more likely to try novel or unconventional products suggested by chatbots. In contrast, conscientiousness and neuroticism did not show significant effects on purchase intention. One explanation is that conscientious consumers prefer transparent and structured decision-making processes and may resist opaque or dynamic algorithmic recommendations (Nordheim et al., 2019). Meanwhile, neurotic consumers often exhibit anxiety and mistrust toward unfamiliar technologies, including AI agents, leading to lower adoption intentions (Wang C. et al., 2023; Wang L. et al., 2023; Wang X. et al., 2023).

Personal innovativeness, defined as the degree to which individuals are open to new technologies (Agarwal and Prasad, 1999), emerged as a strong predictor of chatbot-driven purchases. Highly innovative consumers tend to explore and embrace new tools such as chatbots (Ali and Warraich, 2023). Moreover, personal innovativeness positively moderates the effect of extraversion on purchase intention, suggesting that highly innovative and extraverted consumers engage more frequently and meaningfully with chatbots, increasing their likelihood of purchasing.

The characteristics of product recommendation chatbots also play a crucial role. Expertise, interactivity, and customization significantly influence purchase intention. Consumers are more inclined to trust recommendations from chatbots that demonstrate professional knowledge, offer human-like interaction, and tailor content based on preferences and behavior (Nordheim et al., 2019; Shin and Choi, 2021; Orden-Mejía et al., 2023; Periaiya and Nandukrishna, 2023). Customization, in particular, allows users to feel greater control over the shopping experience, fostering engagement and purchase intent. Conversely, trustworthiness did not significantly affect purchase intention. This finding deviates from traditional source trustworthiness theory but aligns with newer perspectives indicating that Generation Z emphasizes functional benefits—such as speed, usability, and fit—over institutional trust (Reinikainen et al., 2020; Lajante et al., 2023).

Artificial neural network (ANN) analysis provided further insights. While SEM identified personal innovativeness as the most influential factor and customization as third, ANN results indicated that customization had the strongest predictive power, followed by personal innovativeness. This divergence underscores the importance of modeling nonlinear relationships. ANN's capacity to capture complex interactions reveals that consumers assign greater value to highly personalized shopping experiences in real-world settings (Wang C. et al., 2023; Wang L. et al., 2023; Wang X. et al., 2023; Chen and Wu, 2024). Thus, ANN complements the linear assumptions of SEM by providing a more nuanced understanding of consumer behavior.

Lastly, necessary condition analysis (NCA) revealed that agreeableness, openness, expertise, interactivity, and customization are essential for purchase intention—confirming their roles across all analytical methods. Interestingly, extraversion and personal innovativeness, though influential, were not necessary conditions. While extraverted consumers may be more socially inclined, introverts—who prefer solitude or small-group communication—can still exhibit strong purchase intent when supported by personalized chatbot recommendations (Lim et al., 2019). Similarly, although personal innovativeness enhances adoption, modern chatbot systems can accommodate a wide range of user preferences, diminishing the necessity of high innovativeness for generating purchase behavior (Jackson et al., 2013).

Together, these findings highlight the multifaceted psychological and technological determinants of Generation Z's purchase decisions in chatbot-assisted e-commerce environments. By integrating personality psychology with AI-enabled commerce, this study contributes both theoretical clarity and practical implications for chatbot design and personalization strategies.

7 Conclusion

7.1 Theoretical contributions

This study offers several key theoretical contributions to the existing literature on consumer behavior in AI-mediated e-commerce environments, particularly in the context of Generation Z's interactions with product recommendation chatbots. By integrating the Big Five personality traits with the attributes of product recommendation

TABLE 10 Sensitivity analysis.

ANN	Input: EXT, AGR, OPE, EXP, INT, CUS, PEI Output: PI						
	EXT	AGR	CUS	OPE	EXP	INT	PEI
ANN1	0.313	0.032	0.192	0.055	0.035	0.154	0.221
ANN2	0.298	0.054	0.316	0.134	0.035	0.005	0.159
ANN3	0.227	0.194	0.141	0.039	0.148	0.022	0.230
ANN4	0.190	0.255	0.163	0.092	0.034	0.029	0.237
ANN5	0.180	0.010	0.299	0.008	0.149	0.169	0.186
ANN6	0.166	0.028	0.233	0.059	0.137	0.138	0.240
ANN7	0.027	0.259	0.196	0.268	0.078	0.061	0.112
ANN8	0.057	0.286	0.221	0.160	0.006	0.045	0.225
ANN9	0.254	0.171	0.201	0.071	0.008	0.090	0.206
ANN10	0.112	0.037	0.210	0.036	0.204	0.082	0.319
ARI	0.182	0.133	0.217	0.092	0.083	0.080	0.213
NI(%)	83.87%	61.29%	100.00%	42.40%	38.25%	36.87%	98.16%

EXP, Expertise; PEI, Personal Innovativeness; INT, Interactivity; EXT, Extraversion; CUS, Customization; AGR, Agreeableness; OPE, Openness; PI, Purchase Intention; ARI, ANN Relative importance; NI, Normalized Importance.

TABLE 11 Comparison between PLS-SEM and ANN results.

PLS path	Original sample (O)/path coefficient	ANN results: (average relative importance)	Ranking (PLS-SEM) (based on path coefficient)	Ranking (ANN) (based on Average relative importance)	Remark
Input: EXT, AGR, OPE, EXP, INT, CUS, PEI Output: PI					
EXT → PI	0.216	0.182	2	3	Not match
AGR → PI	0.175	0.133	4	4	
OPE → PI	0.149	0.092	5	5	
EXP → PI	0.123	0.083	6	6	
INT → PI	0.097	0.080	7	7	
CUS → PI	0.194	0.217	3	1	
PEI → PI	0.329	0.213	1	2	

EXP, Expertise; PEI, Personal Innovativeness; INT, Interactivity; EXT, Extraversion; CUS, Customization; AGR, Agreeableness; OPE, Openness; PI, Purchase Intention; ARI, ANN Relative importance; NI, Normalized Importance.

chatbots, this research provides a more comprehensive understanding of how personality traits and chatbot characteristics jointly influence purchasing decisions. Our findings contribute to the theoretical understanding of both personality psychology and digital consumer behavior, bridging gaps in previous research and offering new insights into the complex mechanisms that drive Generation Z's responses to chatbot-based product recommendations.

One of the primary contributions of this study is its exploration of the intersection between consumer personality traits and chatbot attributes, which has not been sufficiently examined in prior research. While existing studies have addressed the role of personality in technology adoption (Duong, 2021; Liu et al., 2024) and the influence of chatbot features on user acceptance (Zarouali et al., 2018), little attention has been given to how these factors interact. Our study integrates these two perspectives, examining how the Big Five personality traits—extraversion, agreeableness, conscientiousness, neuroticism, and openness—interact with key features of chatbots, such as expertise, trustworthiness, interactivity, and customization, to shape Generation Z's purchase intentions. This integration bridges an important gap in the literature by offering a multi-faceted model that explains how consumers' psychological traits and the functional features of chatbots jointly influence purchasing behavior. For example, while extraverted individuals are more likely to engage with interactive chatbots and derive greater benefit from personalized recommendations, conscientious individuals tend to prefer chatbots that offer structured and reliable information (Huang et al., 2024; Moisescu et al., 2025). This nuanced understanding is a significant departure from prior research, which often treats chatbot features and personality traits as independent predictors.

Another significant contribution of this study is its empirical investigation of how individual personality traits, particularly extraversion, influence Generation Z's responses to product recommendations. While the role of personality traits in general technology adoption has been well established (Svendsen et al., 2013), this study goes a step further by demonstrating that certain personality traits, such as extraversion, play a particularly strong role in shaping consumer behavior in the context of AI-mediated interactions. Extraverted individuals are more likely to engage with chatbots, appreciating their interactive and socially engaging nature, which in turn increases their purchase intentions. This finding is consistent with prior research that suggests that extraverted individuals enjoy social interaction and seek external sources of stimulation (Sowmya et al., 2023). By linking personality psychology with consumer behavior theory, this study introduces a novel framework that not only extends existing research on technology adoption but also provides a new understanding of how individual personality differences influence purchasing decisions in AI-driven environments.

Furthermore, this research contributes to the literature on chatbot functionality by identifying which specific features of chatbots are most influential in shaping consumer behavior. While previous studies have generally acknowledged the importance of chatbot attributes such as expertise, trustworthiness, and interactivity (Gursoy et al., 2022), they often fail to specify which attributes are most effective in driving user acceptance and purchase intentions. Our findings reveal that expertise, interactivity, and customization significantly affect Generation Z's purchasing behavior, providing a more granular understanding of how these features contribute to chatbot effectiveness. These results extend the current literature by pinpointing the specific functional characteristics that enhance chatbot performance, offering practical implications for both academics and practitioners. For example, personalized chatbot recommendations that align with user preferences are found to be particularly influential in increasing purchase intention among Generation Z consumers, especially those high in openness (Yang et al., 2023).

In addition, the introduction of personal innovativeness as a moderating variable offers a unique contribution to the understanding of consumer behavior in AI-mediated environments. While personal innovativeness has been studied in the context of general technology adoption (Agarwal and Prasad, 1999), it has not been extensively explored in relation to chatbot interactions. By showing that personal innovativeness moderates the relationship between personality traits and purchase intention, this study provides a more dynamic model of consumer behavior, highlighting that consumers who are more open to new technologies are more likely to respond positively to chatbot recommendations. This insight provides a deeper understanding of the variability in user responses to AI-driven interactions and helps to refine the theoretical models of technology acceptance by incorporating individual differences in openness to innovation.

Lastly, the methodological approach used in this study is also a significant contribution to the field. By combining Structural Equation Modeling (SEM), Artificial Neural Networks (ANN), and Necessary Condition Analysis (NCA), this research offers a hybrid approach that captures both linear and nonlinear relationships in consumer decision-making. SEM identifies significant predictors based on linear relationships, while ANN allows for the examination of nonlinear associations and ranks variables according to their predictive power. NCA further identifies essential conditions for the occurrence of purchase behavior, adding an additional layer of understanding to the decision-making process. This integrated methodology enhances both the explanatory power and robustness of the findings, providing a comprehensive framework for future research on consumer behavior in AI-driven environments. It also offers a more refined approach to understanding the complex interactions that drive purchasing decisions in digital commerce (Chen and Wu, 2024; Zhu et al., 2025).

7.2 Practical implications

The findings of this study hold practical implications for e-commerce platforms aiming to influence Generation Z consumers' purchasing behaviors through their interactions with product recommendation chatbots. First, chatbot developers should prioritize emphasizing the personalized benefits of chatbot technology to attract young consumers. Specifically, product recommendation chatbots should highlight their ability to deliver tailored recommendations based on users' past behaviors, preferences, and feedback. To this end, it is essential for managers to continuously upgrade and fine-tune the natural language processing and machine learning capabilities of these systems to ensure they effectively address individual needs.

Second, the study underscores the importance of considering individual differences when designing chatbot-based marketing strategies for younger consumers. A nuanced understanding of users' big five personality traits can enable managers to craft more resonant and effective marketing content. Prior to developing targeted strategies, it is advisable for managers to conduct detailed assessments of personality profiles within their intended user base. Aligning product recommendations with these personality dimensions can enhance the perceived relevance of the offerings, thereby improving user satisfaction and increasing purchase intention.

Moreover, enhancing the expertise and interactivity of chatbots should be a key managerial priority. Product recommendations that demonstrate domain-specific expertise—such as detailed knowledge of product features, user reviews, and current market trends—can bolster user confidence in the chatbot's reliability. Additionally, chatbots that recall users' prior interactions and preferences reflect a deeper understanding of their individual journeys. By engaging users in a friendly, courteous, and human-like manner, chatbots can evolve from functional tools into relatable digital companions, increasing user willingness to interact.

Finally, to appeal to consumers with high personal innovativeness, managers should ensure that product recommendation chatbots embody a degree of technological novelty. Innovative interaction designs—such as immersive interfaces and novel features—can enhance user engagement. Integration of emerging technologies such as voice recognition and virtual reality should be considered to elevate the user experience and stimulate curiosity. In addition, digital platform managers should explore new features that cater to diverse user needs and enhance enjoyment, such as enabling real-time social interaction or gamified recommendation environments.

7.3 Limitation and future research

While this study offers theoretical insights and practical implications, it is subject to several limitations that should be acknowledged. Firstly, the sample is geographically and demographically restricted, focusing solely on generation Z consumers in China. Within this group, the majority of respondents were male (73.33%) and students (72.71%), introducing a gender and student bias that may limit the generalizability of the findings to broader populations. This sample composition may have influenced certain trait-based patterns (e.g., extraversion or openness) and response tendencies, which could differ in more gender-balanced or occupationally diverse populations.

Secondly, the reliance on self-reported questionnaire data may only capture participants' subjective perceptions and intended behaviors, rather than actual purchasing behavior. Although constructs like purchase intention and chatbot experience were well operationalized, the study does not account for potential discrepancies between intention and behavior, especially in real-life decision-making environments.

Thirdly, the cross-sectional nature of the data collection limits our ability to observe temporal changes in consumer behavior. For instance, the influence of personalization or trustworthiness on purchase intention may evolve as users become more familiar with chatbots or as chatbot technology advances.

To address these limitations, future studies could diversify their sample demographics, including participants from different age brackets, occupations, and cultural contexts, to enable comparative cross-cultural analysis. Furthermore, incorporating experimental or neuroscience methods such as EEG or fMRI could provide deeper insight into the cognitive and affective processes underlying generation Z's responses to chatbot interactions. Finally, longitudinal research designs would allow scholars to explore how repeated exposure to chatbots and evolving user preferences affect the dynamics of purchase behavior over time.

Data availability statement

The original contributions presented in the study are publicly available. This data can be found here: https://www.jianguoyun.com/p/DUN6uAwQ_t_QDRiz9oMGIAA. Further inquiries can be directed to the corresponding author.

Ethics statement

The studies involving humans were approved by School of Psychology, Shanghai University of Sport, Shanghai, China. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

JL: Conceptualization, Data curation, Methodology, Writing - original draft. JC: Conceptualization, Data curation, Investigation, Methodology, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that no financial support was received for the research and/or publication of this article.

Acknowledgments

We wish to acknowledge the participants.

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.

Generative AI statement

The authors declare that no Gen AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

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/fpsyg.2025.1454197/full#supplementary-material

References

Agarwal, R., and Prasad, J. (1999). Are individual differences germane to the acceptance of new information technologies? *Decis. Sci.* 30, 361–391. doi: 10.1111/j.1540-5915.1999.tb01614.x

Ali, I., and Warraich, N. F. (2023). Impact of personal innovativeness, perceived smartphone ease of use and mobile self-efficacy on smartphone-based personal information management practices. *Electron. Libr.* 41, 419–437. doi: 10.1108/EL-12-2022-0262

Alkawsi, G., Ali, N., and Baashar, Y. (2021). The moderating role of personal innovativeness and users experience in accepting the smart meter technology. *Appl. Sci.* 11:3297. doi: 10.3390/app11083297

Anbalagan, S. K., Nei, X., Mohan, U., Kanamarlapudi, V. K., Kommalapati, A., and Zhao, X. (2025). "Bridging UI design and Chatbot interactions: applying form-based principles to conversational agents," in *International conference on human-computer interaction*. Cham, Switzerland: Springer Nature Switzerland, 223–231.

Arpaci, I., Karatas, K., Kusci, I., and Al-Emran, M. (2022). Understanding the social sustainability of the Metaverse by integrating UTAUT2 and big five personality traits: a hybrid SEM-ANN approach. *Technol. Soc.* 71:102120. doi: 10.1016/j.techsoc.2022.102120

Arpaci, I., and Kocadag Unver, T. (2020). Moderating role of gender in the relationship between big five personality traits and smartphone addiction. *Psychiatry Q.* 91, 577–585. doi: 10.1007/s11126-020-09718-5

Asaoka, Y., Won, M., Morita, T., Ishikawa, E., and Goto, Y. (2020). Heightened negative affects associated with neurotic personality in behavioral addiction. *Front. Psych.* 11:561713. doi: 10.3389/fpsyt.2020.561713

Attigeri, G., Agrawal, A., and Kolekar, S. V. (2024). Advanced NLP models for technical university information chatbots: development and comparative analysis. *IEEE Access* 12, 29633–29647. doi: 10.1109/ACCESS.2024.3368382

Baumert, A., Schmitt, M., Perugini, M., Johnson, W., Blum, G., Borkenau, P., et al. (2017). Integrating personality structure, personality process, and personality development. *Eur. J. Personal.* 31, 503–528. doi: 10.1002/per.2115

Benet-Martínez, V., and John, O. P. (1998). Los cinco grandes across cultures and ethnic groups: multitrait-multimethod analyses of the big five in spanish and English. *J. Pers. Soc. Psychol.* 75, 729–750. doi: 10.1037/0022-3514.75.3.729

Bleidorn, W., Hopwood, C. J., and Lucas, R. E. (2016). Life events and personality trait change. *J. Pers.* 86, 83–96. doi: 10.1111/jopy.12286

Boudreaux, M. J. (2016). Personality-related problems and the five-factor model of personality. *Personal. Disord. Theory Res. Treat.* 7, 372–383. doi: 10.1037/per0000185

Buecker, S., Maes, M., Denissen, J. J. A., and Luhmann, M. (2020). Loneliness and the big five personality traits: a meta–analysis. *Eur. J. Personal.* 34, 8–28. doi: 10.1002/per.2229

Buelow, M. T., and Cayton, C. (2020). Relationships between the big five personality characteristics and performance on behavioral decision making tasks. *Pers. Individ. Differ.* 160:109931. doi: 10.1016/j.paid.2020.109931

Bunea, O. I., Corboş, R. A., Mişu, S. I., Triculescu, M., and Trifu, A. (2024). The next-generation shopper: a study of generation-Z perceptions of AI in online shopping. *J. Theor. Appl. Electron. Commer. Res.* 19, 2605–2629. doi: 10.3390/jtaer19040125

Busseri, M. A., and Erb, E. M. (2024). The happy personality revisited: re-examining associations between big five personality traits and subjective well-being using meta-analytic structural equation modeling. *J. Pers.* 92, 968–984. doi: 10.1111/jopy.12862

Chang, Y. P., and Zhu, D. H. (2012). The role of perceived social capital and flow experience in building users' continuance intention to social networking sites in China. *Comput. Human Behav.* 28, 995–1001. doi: 10.1016/j.chb.2012.01.001

Chen, J., and Wu, Y. (2024). Would you be willing to purchase virtual gifts during esports live streams? Streamer characteristics and cultural traits. *Comput. Human Behav.* 152:108075. doi: 10.1016/j.chb.2023.108075

Chen, Q., Yin, C., and Gong, Y. (2025). Would an AI chatbot persuade you: an empirical answer from the elaboration likelihood model. *Inf. Technol. People* 38, 937–962. doi: 10.1108/ITP-10-2021-0764

Cuartero, N., and Tur, A. M. (2021). Emotional intelligence, resilience and personality traits neuroticism and extraversion: predictive capacity in perceived academic efficacy. *Nurse Educ. Today* 102:104933. doi: 10.1016/j.nedt.2021.104933

Dawes, J. (2008). Do data characteristics change according to the number of scale points used? An experiment using 5-point, 7-point and 10-point scales. *Int. J. Market. Res.* 50, 61–104. doi: 10.1177/147078530805000106

Denoeux, T., and Lengellé, R. (1993). Initializing backpropagation networks with prototypes. *Neural Netw.* 6, 351–363. doi: 10.1016/0893-6080(93)90003-F

Ding, N., Chen, M., and Hu, L. (2025). Examining the impact of big five personality traits on generation Z designers' subscription to paid AI drawing tools using SEM and FsQCA. *Sci. Rep.* 15:17587. doi: 10.1038/s41598-025-02926-8

Dong, W., Wang, Y., and Qin, J. (2023). An empirical study on impulse consumption intention of livestreaming e-commerce: the mediating effect of flow experience and the moderating effect of time pressure. *Front. Psychol.* 13:1019024. doi: 10.3389/fpsyg.2022.1019024

Dul, J., Van der Laan, E., and Kuik, R. (2020). A statistical significance test for necessary condition analysis. *Organ. Res. Methods* 23, 385–395. doi: 10.1177/1094428118795272

Duong, C. D. (2021). Big five personality traits and green consumption: bridging the attitude-intention-behavior gap. *Asia Pac. J. Mark. Logist.* 34, 1123–1144. doi: 10.1108/APJML-04-2021-0276

Eid, M., Riemann, R., Angleitner, A., and Borkenau, P. (2003). Sociability and positive emotionality: genetic and environmental contributions to the covariation between different facets of extraversion. *J. Pers.* 71, 319–346. doi: 10.1111/1467-6494.7103003

Erdogan, B. Z. (1999). Celebrity endorsement: a literature review. *J. Mark. Manag.* 15, 291–314. doi: 10.1362/026725799784870379

Fajkowska, M. (2017). Personality traits: hierarchically organized systems. *J. Pers.* 86, 36–54. doi: 10.1111/jopy.12314

Fazli-Salehi, R., Torres, I. M., Madadi, R., and Zúñiga, M. Á. (2021). The impact of interpersonal traits (extraversion and agreeableness) on consumers' self-brand connection and communal-brand connection with anthropomorphized brands. *J. Brand Manag.* 29, 13–34. doi: 10.1057/s41262-021-00251-9

Finstad, K. (2010). Response interpolation and scale sensitivity: evidence against 5-point scales. *J. Usability Stud.* 5, 104–110.

Fleischmann, E., Dalkner, N., Fellendorf, F. T., Bengesser, S. A., Lenger, M., Birner, A., et al. (2023). The big five as predictors of cognitive function in individuals with bipolar disorder. *Brain Sci.* 13:773. doi: 10.3390/brainsci13050773

Fornell, C., and Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *J. Mark. Res.* 18, 39–50.

Gan, C. (2016). An empirical analysis of factors influencing continuance intention of mobile instant messaging in China. *Inf. Dev.* 32, 1109–1119. doi: 10.1177/0266666915591298

Gil de Zúñiga, H., Diehl, T., Huber, B., and Liu, J. (2017). Personality traits and social media use in 20 countries: how personality relates to frequency of social media use, social media news use, and social media use for social interaction. *Cyberpsychol. Behav. Soc. Netw.* 20, 540–552. doi: 10.1089/cyber.2017.0295

Gorham, J. (1988). The relationship between verbal teacher immediacy behaviors and student learning. *Commun. Educ.* 37, 40–53. doi: 10.1080/03634528809378702

Gursoy, D., Malodia, S., and Dhir, A. (2022). The metaverse in the hospitality and tourism industry: An overview of current trends and future research directions. *J. Hosp. Mark. Manag.* 31, 527–534.

Han, M. C. (2021). The impact of anthropomorphism on consumers' purchase decision in chatbot commerce. *J. Internet Commer.* 20, 46–65. doi: 10.1080/15332861.2020.1863022

Haslam, N., Bastian, B., and Bissett, M. (2004). Essentialist beliefs about personality and their implications. *Personal. Soc. Psychol. Bull.* 30, 1661–1673. doi: 10.1177/0146167204271182

Haugeland, I. K. F., Følstad, A., Taylor, C., and Bjørkli, C. A. (2022). Understanding the user experience of customer service chatbots: an experimental study of chatbot interaction design. *Int. J. Hum.-Comput. Stud.* 161:102788. doi: 10.1016/j.ijhcs.2022.102788

He, Y., Li, W., and Xue, J. (2022). What and how driving consumer engagement and purchase intention in officer live streaming? A two-factor theory perspective. *Electron. Commerce Res. Appl.* 56:101223. doi: 10.1016/j.elerap.2022.101223

Hopwood, C. J. (2025). Personality functioning, problems in living, and personality traits. *J. Pers. Assess.* 107, 143–158. doi: 10.1080/00223891.2024.2345880

Hörz-Sagstetter, S., Ohse, L., and Kampe, L. (2021). Three dimensional approaches to personality disorders: a review on personality functioning, personality structure, and personality organization. *Curr. Psychiatry Rep.* 23:45. doi: 10.1007/s11920-021-01250-y

Huang, S. C., Silalahi, A. D. K., Eunike, I. J., and Riantama, D. (2024). Understanding impulse buying in E-commerce: the big five traits perspective and moderating effect of time pressure and emotions. *Telemat. Inform. Rep.* 15:100157. doi: 10.1016/j.teler.2024.100157

Huete-Alcocer, N., Martinez-Ruiz, M. P., López-Ruiz, V. R., and Izquiedo-Yusta, A. (2019). Archeological tourist destination image formation: influence of information sources on the cognitive, affective and unique image. *Front. Psychol.* 10:2382. doi: 10.3389/fpsyg.2019.02382

Jackson, J. D., Yi, M. Y., and Park, J. S. (2013). An empirical test of three mediation models for the relationship between personal innovativeness and user acceptance of technology. *Inf. Manag.* 50, 154–161. doi: 10.1016/j.im.2013.02.006

Jiang, Y., Yang, X., and Zheng, T. (2023). Make chatbots more adaptive: dual pathways linking human-like cues and tailored response to trust in interactions with chatbots. *Comput. Human Behav.* 138:107485. doi: 10.1016/j.chb.2022.107485

Jin, E., and Eastin, M. S. (2023). Birds of a feather flock together: matched personality effects of product recommendation chatbots and users. *J. Res. Interact. Mark.* 17, 416–433. doi: 10.1108/JRIM-03-2022-0089

Kelman, H. C. (2017). "Processes of opinion change" in *Attitude change* (London, United Kingdom: Routledge), 205–233.

Kim, S., and Priluck, R. (2025). Consumer responses to generative AI chatbots versus search engines for product evaluation. *J. Theor. Appl. Electron. Commer. Res.* 20:93. doi: 10.3390/jtaer20020093

Kovbasiuk, A., Triantoro, T., Przegalińska, A., Sowa, K., Ciechanowski, L., and Gloor, P. (2025). The personality profile of early generative AI adopters: a big five perspective. *Cent. Eur. Manag. J.* 33, 252–264. doi: 10.1108/CEMJ-02-2024-0067

Kumar, A., Shankar, A., and Nayal, P. (2024). Metaverse is not my cup of tea! An investigation into how personality traits shape metaverse usage intentions. *J. Retail. Consum. Serv.* 77:103639. doi: 10.1016/j.jretconser.2023.103639

Lajante, M., Tojib, D., and Ho, T. I. (2023). When interacting with a service robot is (not) satisfying: the role of customers' need for social sharing of emotion. *Comput. Hum. Behav.* 146:107792. doi: 10.1016/j.chb.2023.107792

Lau, A. J., Tan, G. W. H., Loh, X. M., Leong, L.-Y., Lee, V.-H., and Ooi, K.-B. (2021). On the way: hailing a taxi with a smartphone? A hybrid SEM-neural network approach. *Mach. Learn. Appl.* 4:100034. doi: 10.1016/j.mlwa.2021.100034

Lee, S. A., and Bottomley, J. S. (2021). Personality and grief-related panic attacks: analysis of neuroticism, worry, and anxiety sensitivity. *OMEGA – J. Death. Dying* 86, 849–861. doi: 10.1177/0030222821992192

Lew, S., Tan, G. W. H., Loh, X. M., Hew, J. J., and Ooi, K. B. (2020). The disruptive mobile wallet in the hospitality industry: an extended mobile technology acceptance model. *Technol. Soc.* 63:101430. doi: 10.1016/j.techsoc.2020.101430

Lim, E., Gallagher, E. A., Bui, C. N., Beaulieu, C., Simpson, E., Driscoll, L., et al. (2019). What does my personality says about my friendships in later life? *Innov. Aging* 3:S164. doi: 10.1093/geroni/igz038.586

Li, M., and Wang, R. (2023). Chatbots in e-commerce: the effect of chatbot language style on customers' continuance usage intention and attitude toward brand. *J. Retail. Consum. Serv.* 71:103209. doi: 10.1016/j.jretconser.2022.103209

Lin, Y.-T., Hung, T.-W., and Huang, L. T.-L. (2020). Engineering equity: how AI can help reduce the harm of implicit bias. *Philos. Technol.* 34, 65–90. doi: 10.1007/s13347-020-00406-7

Liu, A. Y., Botella Carrubi, M. D., and Blanco González-Tejero, C. (2024). Community startup businesses: the impact of big five personality traits and social media technology acceptance on group buying leaders. *Int. J. Entrep. Behav. Res.* 30, 955–978. doi: 10.1108/IJEBR-07-2023-0685

Liu, D., and Campbell, W. K. (2017). The big five personality traits, big two metatraits and social media: a meta-analysis. *J. Res. Pers.* 70, 229–240. doi: 10.1016/j.jrp.2017.08.004

Liu, X., Wang, D., Gu, M., and Yang, J. (2023). Research on the influence mechanism of anchors' professionalism on consumers' impulse buying intention in the livestream shopping scenario. *Enterp. Inf. Syst.* 17:2065457. doi: 10.1080/17517575.2022.2065457

Li, Y., and Shin, H. (2023). Should a luxury brand's chatbot use emoticons? Impact on brand status. *J. Consum. Behav.* 22, 569–581. doi: 10.1002/cb.2104

Lo, P. S., Dwivedi, Y. K., Tan, W. H., Ooi, K.-B., Aw, E. C.-X., and Metri, B. (2022). Why do consumers buy impulsively during live streaming? A deep learning-based dual-stage SEM-ANN analysis. *J. Bus. Res.* 147, 325–337. doi: 10.1016/j.jbusres.2022.04.013

Louwen, C., Reidlinger, D., and Milne, N. (2023). Profiling health professionals' personality traits, behaviour styles and emotional intelligence: a systematic review. *BMC Med. Educ.* 23:120. doi: 10.1186/s12909-023-04003-y

Lu, Z., Min, Q., Jiang, L., and Chen, Q. (2024). The effect of the anthropomorphic design of chatbots on customer switching intention when the chatbot service fails: an expectation perspective. *Int. J. Inf. Manag.* 76:102767. doi: 10.1016/j.ijinfomgt.2024.102767

Lynn, M. (2021). Effects of the big five personality traits on tipping attitudes, motives, and behaviors. *Int. J. Hospit. Manag.* 92:102722. doi: 10.1016/j.ijhm.2020.102722

Marengo, D., Sindermann, C., Elhai, J. D., and Montag, C. (2020). One social media company to rule them all: associations between use of Facebook-owned social media platforms, sociodemographic characteristics, and the big five personality traits. *Front. Psychol.* 11:936. doi: 10.3389/fpsyg.2020.00936

Meier, M., Maier, C., Thatcher, J. B., and Weitzel, T. (2024). Chatbot interactions: how consumption values and disruptive situations influence customers' willingness to interact. *Inf. Syst. J.* 34, 1579–1625. doi: 10.1111/isj.12507

Moisescu, O. I., Gică, O. A., Herle, F. A., Dan, I., and Sarstedt, M. (2025). Does one size fit all? The role of extraversion in generating electronic word-of-mouth through social media brand page engagement. *Psychol. Mark.* 42, 1827–1847. doi: 10.1002/mar.22205

Neter, J., Wasserman, W., and Kutner, M. H. (1990). Polynomial regression. *Neter J, Wasserman W, Kutner MH. Applied linear statistical models*. Boston: Irwin, 315–41.

Nordheim, C. B., Følstad, A., and Bjørkli, C. A. (2019). An initial model of trust in chatbots for customer service—findings from a questionnaire study. *Interact. Comput.* 31, 317–335. doi: 10.1093/iwc/iwz022

Ohanian, R. (1991). The impact of celebrity spokesperson's perceived image on consumers' intention to purchase. *J. Advert. Res.* 31, 46–52.

Orden-Mejía, M., Carvache-Franco, M., Huertas, A., Carvache-Franco, O., and Carvache-Franco, W. (2023). Modeling users' satisfaction and visit intention using AI-based chatbots In A. Luqman (Ed.),. *PLoS One* 18:e0286427. doi: 10.1371/journal.pone.0286427

Pavlou, P. A. (2003). Consumer acceptance of electronic commerce: integrating trust and risk with the technology acceptance model. *Int. J. Electron. Commer.* 7, 101–134. doi: 10.1080/10864415.2003.11044275

Peng, X., Zhou, L., Tang, Q., and Liu, Y. (2023). "Effects of the conversation and recommendation mechanism on Chatbots' recommendation effectiveness" in *International conference on human-computer interaction* (Cham: Springer Nature Switzerland), 493–501.

Periaiya, S., and Nandukrishna, A. T. (2023). What drives user stickiness and satisfaction in OTT video streaming platforms? A mixed-method exploration. *Int. J. Hum.-Comput. Interact.* 40, 2326–2342. doi: 10.1080/10447318.2022.2160224

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, 879–903. doi: 10.1037/0021-9010.88.5.879

Preston, C. C., and Colman, A. M. (2000). Optimal number of response categories in rating scales: reliability, validity, discriminating power, and respondent preferences. *Acta Psychol.* 104, 1–15. doi: 10.1016/S0001-6918(99)00050-5

Rahevar, M., and Darji, S. (2024). The adoption of AI-driven chatbots into a recommendation for e-commerce systems to targeted customer in the selection of product. *Int. J. Manag. Econ. Commer.* 1, 128–137. doi: 10.62737/m1vpdq75

Reinikainen, H., Kari, J. T., and Luoma-Aho, V. (2020). Generation Z and organizational listening on social media. *Media Commun.* 8, 185–196. doi: 10.17645/mac.v8i2.2772

Richter, N. F., Schubring, S., Hauff, S., Ringle, C. M., and Sarstedt, M. (2020). When predictors of outcomes are necessary: guidelines for the combined use of PLS-SEM and NCA. *Ind. Manag. Data Syst.* 120, 2243–2267. doi: 10.1108/IMDS-11-2019-0638

Riedl, R. (2022). Is trust in artificial intelligence systems related to user personality? Review of empirical evidence and future research directions. *Electron. Mark.* 32, 2021–2051. doi: 10.1007/s12525-022-00594-4

Sadeq, N. A., Valdes, E. G., Harrison Bush, A. L., and Andel, R. (2018). The role of personality in a regular cognitive monitoring program. *Alzheimer Dis. Assoc. Disord.* 32, 226–231. doi: 10.1097/WAD.0000000000000236

Salmony, F. U., and Kanbach, D. K. (2022). Personality trait differences across types of entrepreneurs: a systematic literature review. *Rev. Manag. Sci.* 16, 713–749. doi: 10.1007/s11846-021-00466-9

Sardar, S., Tata, S. V., and Sarkar, S. (2024). Examining the influence of source factors and content characteristics of influencers' post on consumer engagement and purchase intention: a moderated analysis. *J. Retail. Consum. Serv.* 79:103888. doi: 10.1016/j.jretconser.2024.103888

Seyfi, S., Kim, M. J., Nazifi, A., Murdy, S., and Vo-Thanh, T. (2025). Understanding tourist barriers and personality influences in embracing generative AI for travel planning and decision-making. *Int. J. Hospit. Manag.* 126:104105. doi: 10.1016/j.ijhm.2025.104105

Shin, H., Bunosso, I., and Levine, L. R. (2023). The influence of chatbot humour on consumer evaluations of services. *Int. J. Consum. Stud.* 47, 545–562. doi: 10.1111/ijcs.12849

Shin, H.-R., and Choi, J.-G. (2021). The moderating effect of 'generation' on the relations between source credibility of social media contents, hotel brand image, and purchase intention. *Sustainability* 13:9471. doi: 10.3390/su13169471

Skjuve, M., Følstad, A., Fostervold, K. I., and Brandtzaeg, P. B. (2021). My chatbot companion - a study of human-chatbot relationships. *Int. J. Hum.-Comput. Stud.* 149:102601. doi: 10.1016/j.ijhcs.2021.102601

Sowmya, D., Debbata, K., Vignesh, G. S., Sagar, D. S. N., Bavanthula, A. R., and Pallavi, L. (2023). "Emerging role of healthcare chatbots in improving medical assistance," in *2023 3rd Asian Conference on Innovation in Technology (ASIANCON)*. IEEE. p. 1–6.

Stavrova, O., Evans, A. M., and van Beest, I. (2022). The effects of partner extraversion and agreeableness on trust. *Personal. Soc. Psychol. Bull.* 49, 1028–1042. doi: 10.1177/01461672221086768

Svendsen, G. B., Johnsen, J.-A. K., Almås-Sørensen, L., and Vittersø, J. (2013). Personality and technology acceptance: the influence of personality factors on the core constructs of the technology acceptance model. *Behav. Inf. Technol.* 32, 323–334. doi: 10.1080/0144929X.2011.553740

Tsai, W.-H. S., Liu, Y., and Chuan, C.-H. (2021). How chatbots' social presence communication enhances consumer engagement: the mediating role of parasocial interaction and dialogue. *J. Res. Interact. Mark.* 15, 460–482. doi: 10.1108/JRIM-12-2019-0200

Tucaković, L., and Nedeljković, B. (2022). Personality and affective correlates of openness to experience from big five and HEXACO personality models: the dual nature of big five openness. *J. Pers. Assess.* 105, 544–554. doi: 10.1080/00223891.2022.2117047

Venkatesh, V., JYL, T., and Xu, X. (2012). Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology. *MIS Q.* 36:157. doi: 10.2307/41410412

Walley, K., Goodall, S., Humphries, A., Huntington, J., White, D., and Asson, T. (2017). Key dimensions of personal innovativeness. *Int. J. Bus. Innov. Res.* 12:259. doi: 10.1504/IJBIR.2017.081408

Wang, C. L. (2023). "Interactive marketing is the new normal" in *The Palgrave handbook of interactive marketing* (Cham: Springer International Publishing), 1–12.

Wang, C., Li, Y., Fu, W., and Jin, J. (2023). Whether to trust chatbots: applying the event-related approach to understand consumers' emotional experiences in interactions with chatbots in e-commerce. *J. Retail. Consum. Serv.* 73:103325. doi: 10.1016/j.jretconser.2023.103325

Wang, L., Li, X., Zhu, H., and Zhao, Y. (2023). Influencing factors of livestream selling of fresh food based on a push-pull model: a two-stage approach combining structural equation modeling (SEM) and artificial neural network (ANN). *Expert Syst. Appl.* 212:118799. doi: 10.1016/j.eswa.2022.118799

Wang, X., Luo, R., Liu, Y., Chen, P., Tao, Y., and He, Y. (2023). Revealing the complexity of users' intention to adopt healthcare chatbots: a mixed-method analysis of antecedent condition configurations. *Inf. Process. Manag.* 60:103444. doi: 10.1016/j.ipm.2023.103444

Wilmot, M. P., and Ones, D. S. (2022). Agreeableness and its consequences: a quantitative review of meta-analytic findings. *Personal. Soc. Psychol. Rev.* 26, 242–280. doi: 10.1177/10888683211073007

Yang, X., Liu, Y., Dong, J., and Li, S. (2023). Impact of streamers' characteristics on sales performance of search and experience products: evidence from Douyin. *J. Retail. Consum. Serv.* 70:103155. doi: 10.1016/j.jretconser.2022.103155

Yin, J., and Qiu, X. (2021). "Influence mechanism research on the AI Marketing Technology of Online Shopping Platform on consumers' purchase intention" in *6th Annual International Conference on Social Science and Contemporary Humanity Development (SSCHD 2020)* (Dordrecht, The Netherlands: Atlantis Press), 833–838.

Zarouali, B., Van den Broeck, E., Walrave, M., and Poels, K. (2018). Predicting consumer responses to a Chatbot on Facebook. *Cyberpsychol. Behav. Soc. Netw.* 21, 491–497. doi: 10.1089/cyber.2017.0518

Zhong, H., and Han, Z. (2023). A systematic review of information source preference research. *J. Doc.* 80, 255–274. doi: 10.1108/JD-03-2023-0059

Zhu, Y. P., Xin, L., Wang, H., and Park, H. W. (2025). Effects of AI virtual anchors on brand image and loyalty: insights from perceived value theory and SEM-ANN analysis. *Systems* 13:79. doi: 10.3390/systems13020079

Zweig, D., and Webster, J. (2003). Personality as a moderator of monitoring acceptance. *Comput. Human Behav.* 19, 479–493. doi: 10.1016/S0747-5632(02)00075-4

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 08 September 2025
DOI 10.3389/fpsyg.2025.1636906

OPEN ACCESS

EDITED BY
Qixing Qu,
University of International Business and Economics, China

REVIEWED BY
Xicheng Yin,
University of Electronic Science and Technology of China, China
Jian Jin,
Beijing Normal University, China

*CORRESPONDENCE
Ruqin Ren
✉ ruqinren@sjtu.edu.cn

RECEIVED 28 May 2025
ACCEPTED 25 August 2025
PUBLISHED 08 September 2025

CITATION
Han J and Ren R (2025) Why unequal AI access enhances team productivity: the mediating role of interaction processes and cognitive diversity.
Front. Psychol. 16:1636906.
doi: 10.3389/fpsyg.2025.1636906

Why unequal AI access enhances team productivity: the mediating role of interaction processes and cognitive diversity

Jiaxuan Han and Ruqin Ren*

USC-SJTU Institute of Cultural and Creative Industry, Shanghai Jiao Tong University, Shanghai, China

Introduction: Generative artificial intelligence (GenAI) is widely viewed as valuable for improving the performance of human-agent teams (HATs). However, in reality, not all members have equal access to AI tools, making uneven AI integration an important factor impacting team composition and, thus, team effectiveness. While unequal access might seem detrimental, potentially hindering technology utilization, it could also foster deeper interactions and diverse expertise. To clarify these mechanisms, this study extends the classic Input-Mediator-Output model to an Input-Process-State-Output (IPSO) framework.
Methods: A lab experiment involving 60 two-person teams was conducted, with teams assigned to unequal, full, or no AI access conditions.
Results: The findings indicate that unequal AI access yields the highest productivity, improving both task quality and completion time compared to no or full AI access. This effect is driven by two key mechanisms. First, negative socio-emotional interactions and increased cognitive diversity serve as a positive serial mediation pathway linking unequal AI access to enhanced task quality. Second, unequal AI access leads to more concentrated and imbalanced questioning behaviors, which accelerates task completion.
Discussion: This study provides an in-depth theoretical explanation of how AI integration structures operate in HATs and offers a foundation for strategically optimizing GenAI access in human-agent teaming.

KEYWORDS

AI access, cognitive diversity, team interaction processes, team productivity, human-agent teams, human-AI collaboration

1 Introduction

As generative AI (GenAI) technology continues to evolve, more individuals and organizations are integrating GenAI into collaborative work, forming Human-Agent Teams (HATs). HAT refers to a collaborative effort between one or more humans and autonomous agents to achieve a common goal (McNeese et al., 2018; O'Neill et al., 2022). A recent industry report found that 78% surveyed organizations adopted AI in their organizations, with 56% of employees directly engaging with AI tools to automate or augment job tasks (BusinessWire, 2023). Despite GenAI's widespread application, challenges remain—particularly regarding the often complex and inconsistent ways team members adopt AI. It cannot be taken for granted that AI access among team members is equal. In practice, some team members use GenAI extensively, while others lack access or proficiency (Humlum and Vestergaard, 2025), resulting in diverse AI integration structures within Human-Agent Teams.

The challenge of inconsistent AI access is particularly salient in short-term project-based team settings, which are often termed *ad hoc* or temporary teams (Finholt et al., 2014; Liu et al., 2004; Majchrzak et al., 2012). Unlike long-standing corporate teams, people in temporary teams lack prior relationships and must collaborate effectively with minimal knowledge of each other. In these settings, AI tools become important external resources. Moreover, as temporary teams typically lack clearly specified management hierarchies or power structures (Stone et al., 2010), technological asymmetries may carry greater weight in shaping team dynamics. Thus, the uneven distribution of AI access raises important questions about how different GenAI access patterns affect already complex and challenging temporary team collaboration.

Extant literature has demonstrated that AI adoption influences team productivity, which is defined as the collective effectiveness (i.e., task quality) and efficiency (i.e., task time) (Hackman, 1978; Kozlowski and Ilgen, 2006; Kwarteng et al., 2023; Noy and Zhang, 2023). However, how and why GenAI integration structures might influence team productivity remains a subject of theoretical debate. Though it seems intuitive to assume that equipping all members with the most advanced technology would be optimal, given widely existing evidence that GenAI usage increases individual users' creativity and productivity (Cui et al., 2024; Doshi and Hauser, 2024; Noy and Zhang, 2023). Limiting AI access may also result in imbalanced participation and decreased morale and contribution from those without access (Bayerl et al., 2016; Rogers et al., 2009; Simaremare et al., 2024). However, there also exist counterarguments that limiting AI touchpoints may enhance team interactions (Li et al., 2024; Raisch and Krakowski, 2021) and encourage diverse perspectives to emerge as the team could tap into both personal expertise and AI outputs, rather than having all members quickly converging on the same AI-generated outputs (Doshi and Hauser, 2024).

To resolve conflicting views on the optimal strategy for GenAI adoption in HATs, the current study explores how full AI access, partial AI access, and no AI access shapes team dynamics differently, and how these dynamics, in turn, influence collaborative performance. Unequal AI access is of particular interest as it introduces distinct intra-team dynamics that are less likely to emerge in uniformly equipped teams, including asymmetric information distribution (Sebo et al., 2018; Gurkan and Yan, 2023; Zvelebilova et al., 2024), divergent expectations of contribution (Doshi and Hauser, 2024; Stasser and Titus, 2003; Lu et al., 2012), and shifts in perceived social status (Rogers et al., 2009; Meeussen and Van Dijk, 2016). Such dynamics represent novel organizational conditions that may fundamentally reshape how teams interact, adapt, and perform. Despite its increasing relevance, prior research has primarily contrasted teams with full AI access and those without (e.g., Han et al., 2024; Gurkan and Yan, 2023), overlooking this nuanced middle ground. The findings illuminate both the practical implications of AI integration in teamwork and the theoretical significance of how unequal access reshapes team interaction and productivity.

We draw on O'Neill et al. (2023)'s recent extension of the classic Input–Mediator–Output (IMO) model (Hackman, 1978; Ilgen et al., 2005; Marks et al., 2001). The IMO model has historically been used in research on human team effectiveness and small group interactions (Hackman, 1978; Steiner, 1972; Ilgen et al., 2005; Marks et al., 2001), providing a structured lens to examine how team inputs (e.g., member composition, task design) influence team outputs (e.g., performance, satisfaction) via mediating mechanisms such as team processes and emergent states. O'Neill et al. (2023) applied the IMO model HATs, providing a framework for examining how inputs unique to HATs—such as different modes of human-AI composition—shape mediating team dynamics and ultimately affect outcomes. To better adapt this umbrella framework to our research context, we now propose two conceptual modifications to explain how AI integration patterns (input) affect team productivity (output) in greater detail.

First, regarding team *input*, we conceptualized varied AI integration patterns as a key team composition factor. Team composition is the different ways that human-autonomy is combined in HATs. Most existing studies treat AI usage as a binary input—either present or absent (Al Naqbi et al., 2024; Gohar and Utley, 2023; Gurkan and Yan, 2023; Han et al., 2024), overlooking the nuanced AI integration structures that more accurately reflect real-world practices. For example, Han et al. (2024), in their examination of the effects of GenAI on team collaboration in creative tasks, included only two conditions: human teams with GenAI and without GenAI. Similarly, Gurkan and Yan (2023) designed their experiment such that a chatbot provides information in a group chat without engaging in direct interaction, considering only the presence or absence of AI when evaluating its effects on cognitive diversity and team decision-making. Such designs oversimplified the patterns of GenAI allocation among team members. To better capture the nuances of AI adoption in reality, we aim to explore how AI integration structures as a team input impact team processes and outcomes by carefully considering three conditions of AI integration: no access to AI, partial access to AI, and full access to AI among the team.

Second, for the *mediator* part, O'Neill et al. (2023) emphasized the importance of considering mediating mechanisms and moving beyond a simplistic independent-dependent variable modeling approach. In their framework, the mediator was conceptually divided into two broad categories: interaction processes (e.g., planning, communication, coordination) and emergent states (e.g., trust, shared mental models, situation awareness, or affective states). However, they did not specify the potential relationships between these two types of mediators. We further propose a sequential relationship between them: interactions processes, as manifested by individual members' communication behaviors, give rise to emergent states (cognitive or affective) at the team level. In other words, emergent states are not static but dynamically shaped through interactions. Thus, we delineate the mediator part into two consecutive steps and propose them as chained mediators, transforming the *Input-Mediator-Output (IMO)* model into an *Input-Process-States-Output (IPSO)* model, which we then subject to empirical testing. Figure 1 illustrates how we further modify the IMO model for HATs proposed by O'Neill et al. (2023).

For this current study, we focus on communication behaviors as the 'Process' factor and cognitive diversity as the 'State' factor in our IPSO model. While cognitive diversity is a classic construct in teaming research and is often recognized as a team emergent property evolving through dynamic interactions (Marks et al., 2001; Mello and Rentsch, 2015), and some initial HATs research links AI usage to cognitive diversity (Gurkan and Yan, 2023), these studies often stop short of identifying specific interaction behaviors that mediate this relationship. We will use Bales' Interaction Process Analysis (IPA) to classify four specific types of communication behaviors and explore how they potentially alter task-related information flow and contribute to cognitive divergence among members.

In conclusion, our IPSO model aims to provide a more accurate description of how different GenAI access structures influence team interaction patterns and cognitive diversity, and how these factors jointly impact team outcomes such as task quality and completion time. Accordingly, we attempt to address this general question:

How do varied GenAI integration structures affect team productivity via the serial mediation mechanisms of team interaction behaviors and cognitive diversity?

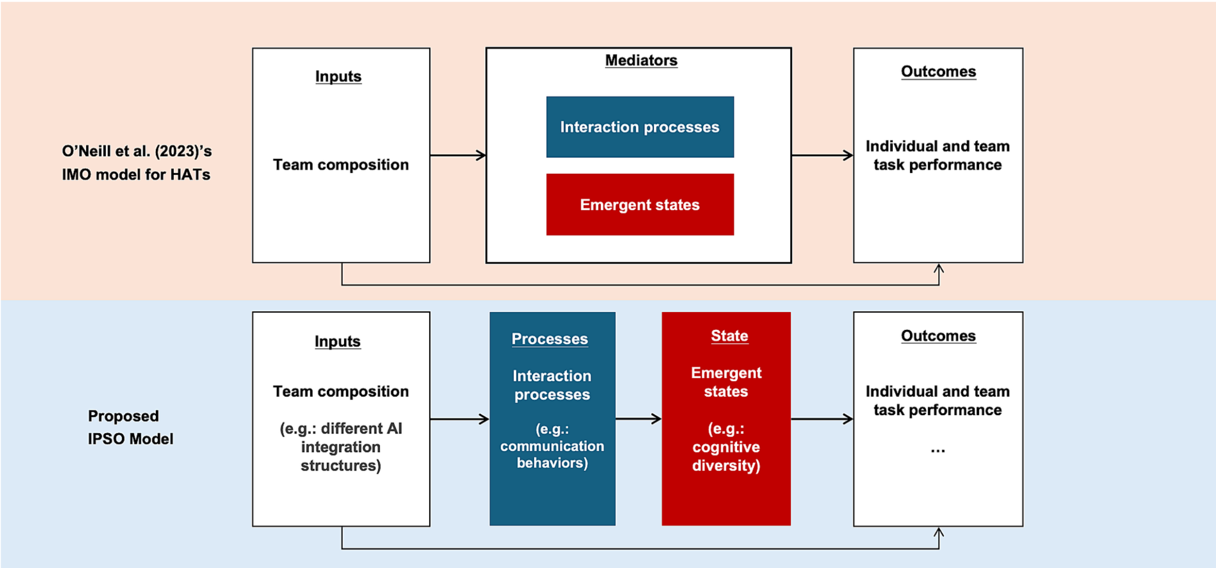


FIGURE 1
IPSO model: an extended framework for O'Neill et al. (2023)'s IMO model for HATs.

2 Hypotheses development

2.1 The paradox of unequal AI integration

Recent studies have consistently shown that integrating AI into human teams can enhance collaborative outcomes by fostering creativity (Jeong and Jeong, 2024), improving decision-making (Gurkan and Yan, 2023), and boosting productivity (Al Naqbi et al., 2024). However, moving beyond this binary perspective of AI adoption, real-world scenarios often involve uneven access to AI within teams. When examining how such unequal distribution of this emerging technology affects team productivity, prior research offers conflicting conclusions. The *positive perspectives* suggest that limiting AI access to some of the team members facilitates more focused and interactive use of GenAI, thus enhancing its utilization depth and maximizing its potential (Raisch and Krakowski, 2021). This close human-AI collaboration can foster creativity and improve task quality (Zhang et al., 2025). Additionally, the selective use of GenAI by only some team members helps to generate diverse cognitive inputs and reduce homogeneous ideas (Doshi and Hauser, 2024), teams thus may achieve high levels of creativity by building a wider pool of expertise that is differentiated and specialized (Zhang et al., 2025). With respect to task time, unequal AI access can shorten completion time by streamlining communication and facilitating strategic adjustments (Li et al., 2022), as full access may increase coordination complexity with many more human-AI pairings to manage (Becker et al., 2008). Limiting such human-AI combinations can reduce communication costs and accelerate task execution. Moreover, the diverse inputs resulting from unequal access to technology can make teams more flexible and agile (Pieterse et al., 2011), allowing them to adjust more quickly in the face of change and unexpected situations (Harrison et al., 2000).

The *opposing viewpoint* suggests that full access to new technology is more beneficial for task quality because it fosters equal participation among team members (Rogers et al., 2009), potentially maximizing each individual's contribution to the team (Li et al., 2024). When technology distribution is not equal, those without AI access may feel marginalized, which can diminish their motivation to participate and contribute actively (Bayerl et al., 2016), ultimately leading to lower overall team cohesion and reduced task quality. Furthermore, uneven AI distribution may prolong task completion time by increasing the difficulty of managing conflict and interpersonal tension caused by unequal participation among team members (Bankins and Formosa, 2023; Rogers et al., 2009). In addition, the diverse perspectives generated by varying collaboration patterns often require more extensive integration efforts to reach a consensus (Sauer et al., 2006), all of which demand additional time (Mohammed and Schillinger, 2022; Narayan et al., 2021). Therefore, we propose a set of competing hypotheses:

> *H1a*: Teams with *unequal* AI access have greater team productivity (i.e., better task quality and faster task completion time) compared to those with no access or full access.
>
> *H1b*: Teams with *full* AI access have better team productivity (i.e., better task quality and faster task completion time) compared to those with no access or unequal access.

2.2 The mediator role of cognitive diversity between AI integration and team productivity

Building on the above discussion, a likely mechanism through which AI integration structures influence team productivity is the diverse task-related perspectives and contributions that stem from differences in access, a concept commonly referred to as cognitive diversity. It is defined as the range of information, information processing styles, and perspectives of members, which is dynamically and interactively generated through communication (Gurkan and Yan, 2023; Sauer et al., 2006). Though cognitive diversity is a complex construct and has been defined in many varied ways (Kurtzberg, 2005; Sauer et al., 2006; Shin et al., 2012; Miller et al., 1998; Mohammed and Ringseis, 2001), Mello and Rentsch (2015) proposed a stability-based framework that categorizes cognitive diversity into four types, ranging from the most stable to the most malleable: trait-like (stable and consistent personal characteristics), developmental (which evolve over time but change gradually), acquired (context-dependent and flexible, such as task-related knowledge or attitudes), and exposed (the most malleable, shaped by specific experimental conditions). Our study specifically focuses on *acquired* cognitive diversity, which evolve dynamically with team context. This form of cognitive diversity is important for understanding team collaboration in our research context, given its direct susceptibility to variation in members' access to external information sources, particularly AI technology and how it is integrated within teams.

We speculate that not distributing AI access equally within teams could lead to increased cognitive diversity mainly by triggering task-related information asymmetry and social status and role differentiation. First, unequal AI integration reshapes how information is accessed and shared within teams, leading to differences in members' task-related information processing and perspectives (DeSanctis and Poole, 1994; Zhang et al., 2025). When only some team members have access to AI assistance while others do not, they are exposed to different sources of task-relevant information. AI-equipped members may form task opinions based on algorithmic interpretations or AI-generated contents (Gurkan and Yan, 2023; Sebo et al., 2018; Zvelebilova et al., 2024), whereas non-AI users rely on human discussions, intuition, or personal experience. In contrast, teams with full AI access could use highly similar information, as members largely depend on the homogenized outputs generated by AI (Doshi and Hauser, 2024). According to social confirmation bias, this shared information often overshadows unique insights derived from individual knowledge or experience (Lu et al., 2012; Stasser and Titus, 2003), easily results in convergent perspectives within the team. Therefore, unequal AI access is likely to foster greater cognitive diversity by generating a wider range of opinions arising from distinct informational environments.

Second, AI access serves as a substitute for human expertise (Doshi and Hauser, 2024; Korzynski et al., 2023; Noy and Zhang, 2023; Zhang et al., 2023), creating role and status differentiation between users and non-users. People with GenAI access may perceive themselves—and be perceived by others—as more competent due to their technological advantage (Meeussen and Van Dijk, 2016; Rogers et al., 2009). Drawing on status characteristics theory (Berger et al., 1980; Correll and Ridgeway, 2003), AI access could serve as a salient status characteristic, shaping interaction patterns and authority structures within teams (Zhang et al., 2025). High-status individuals typically make strategic decisions, while lower-status members focus on operational aspects of the task (Bunderson and Reagans, 2011). Such differentiated roles and statuses—emerging from unequal AI access—further contribute to more varied information processing styles and task-related perspectives among team members (Mello and Rentsch, 2015).

As the critical team-level psychological outcome of unequal AI access, increased cognitive diversity is commonly related to both positive and negative team-level outcomes (Horwitz and Horwitz, 2007; Simons and Rowland, 2011), such as task quality (Gomez and Lazer, 2019; Joniaková et al., 2021; Patrício and Franco, 2022; Schumpe et al., 2023) and task time (Harrison et al., 2000; Li et al., 2022; Mohammed and Schillinger, 2022; Sauer et al., 2006). We, therefore, posit it as a mediator between AI integration and team productivity without predicting directionality. It is hypothesized that:

> *H2*: Cognitive diversity mediates the relationship between AI integration structure and team productivity.

2.3 The mediator role of team interaction processes between AI integration and cognitive diversity

Team interaction, as a dynamic process central to team functioning, plays a critical role in shaping emergent states such as cognitive diversity (Marks et al., 2001; Mello and Rentsch, 2015). Prior sections discussed how unequal AI access may create informational asymmetry and status differentiation within teams. These effects can directly alter how members exchange information and relate to one another (Ward, 2013), thereby affecting both cognitive diversity and productivity. To further unpack team interactions as observable actions, this study adopts Bales' Interaction Process Analysis (IPA) framework (Bales, 1950; Nam et al., 2009; Soukup et al., 2020), a well-established categorizing scheme for team interactions. IPA separates complex team interactions into socio-emotional (positive/negative reactions) and task-related (questions/answers) domains, providing a structured approach for analyzing how different interaction patterns emerge under varying GenAI access conditions.

2.3.1 Unequal AI access's impact on socio-emotional area interactions

Socio-emotional interactions can be further divided into positive and negative reactions. Positive reactions include showing solidarity, releasing tension, and expressing agreement, while negative reactions refer to disagreement, tension, and antagonism (Nam et al., 2009). Unequal AI access can shape these emotional reactions in contrasting ways—potentially suppressing supportive behaviors due to perceived unfairness while at the same time encouraging disagreement as a result of divergent informational inputs (Mannes et al., 2014; Pelled et al., 1999).

First, perceived inequality in the distribution of a highly desirable technology may lead to misunderstanding and mistrust, thereby reducing the expression of positive interactions like support or agreement (Cronin et al., 2011; Kennedy and Pronin, 2008). This undermines the development of a psychologically safe environment that encourages broad participation and open perspective-sharing, ultimately hindering the emergence of cognitive diversity (Isohätälä et al., 2020). For example, repeatedly interrupting others' views during group discussions may trigger defensiveness and discourage the contribution of diverse ideas. Thus,

> *H3a*: Unequal AI access *reduces positive* socio-emotional behaviors, which in turn influence cognitive diversity.

Second, unequal AI integration can increase negative reactions such as disagreement by encouraging the exchange of unique and unshared information from distinct perspectives (Lu et al., 2012)—AI users draw on system-generated content, while non-users rely more on personal experience. When team members challenge one another's assumptions or interpretations, they may surface divergent mental models and expose hidden knowledge structures, which in turn promotes deeper discussion and helps teams avoid premature consensus (Cronin et al., 2011; Mohammed et al., 2023; Srikanth et al., 2016). In this way, negative socio-emotional expressions may reflect more diverged rather than converged communication, contributing to richer team cognition (Mohammed et al., 2023). Thus,

> *H3b*: Unequal AI access *increases negative* socio-emotional behaviors, which in turn influence cognitive diversity.

2.3.2 Unequal AI access' impact on task area interactions

In the IPA framework, task-related interactions are divided into questioning (e.g., asking for suggestions, opinions, or information) and answering behaviors (e.g., providing suggestions, opinions, or information) (Nam et al., 2009). Unequal AI integration creates information asymmetries and initial status expectation differences, as timely information and content-generation capabilities are more readily available to AI users. This results in concentrated questioning and answering behaviors, ultimately influencing cognitive diversity (Bunderson and Reagans, 2011).

First, non-AI users, facing information disadvantages, are more likely to seek orientation or advice from AI-equipped teammates to compensate for knowledge gaps. Simultaneously, AI users, seen as knowledge contributors, tend to take on the role of providing task-relevant input to facilitate team coordination (Rogers et al., 2009; Zhang et al., 2025). As a result, task-related communication—both questioning and answering—becomes increasingly concentrated. This interactional imbalance resulting from informational asymmetry can shape how information flows and integrates into teams, further affecting cognitive diversity. Specifically, such imbalanced information exchanges expose non-overlapping cognitive regions and stimulate cross-boundary information flow, which promotes knowledge integration (Bunderson and Sutcliffe, 2002; Mesmer-Magnus and DeChurch, 2009) and deepens analytical engagement (Homan et al., 2007). Ultimately, such patterns support the emergence of greater team cognitive diversity.

Moreover, access to AI may elevate expectations about one's task contributions, making AI users often perceived as high-status actors within teams (Correll and Ridgeway, 2003). These status differences shape the direction of task-related communication (Bunderson and Reagans, 2011). For instance, higher-status members are more likely to assume directive roles by offering orientations and suggestions, whereas lower-status members tend to ask more questions and seek guidance from those perceived as more knowledgeable (Chung and Pennebaker, 2011; De Jong et al., 2022). Building on this dynamic, unequal AI access may initially create status-based expectations that result in questioning and answering behaviors becoming concentrated within specific individuals. Over time, such interaction patterns can reinforce and solidify team status hierarchies, which represent differentiated perspectives and styles in approaching tasks (Harrison and Klein, 2007). Thus, we propose the following hypotheses:

> *H3c*: Unequal AI access increases concentrated task-related *questioning*, which in turn influences cognitive diversity.
>
> *H3d*: Unequal AI access increases concentrated task-related *answering*, which in turn influences cognitive diversity.

The hypothesized research model, as depicted in Figure 2, integrates the serial mediation links between varied AI access structures, team interaction processes (further divided into socio-emotional and task-oriented processes), cognitive diversity, and team productivity (quality and time).

3 Method

We conducted a randomized and controlled laboratory experiment to examine how different AI integration structures influence team cognitive diversity and task performance through a press release writing task. The study recruited a total of 120 university

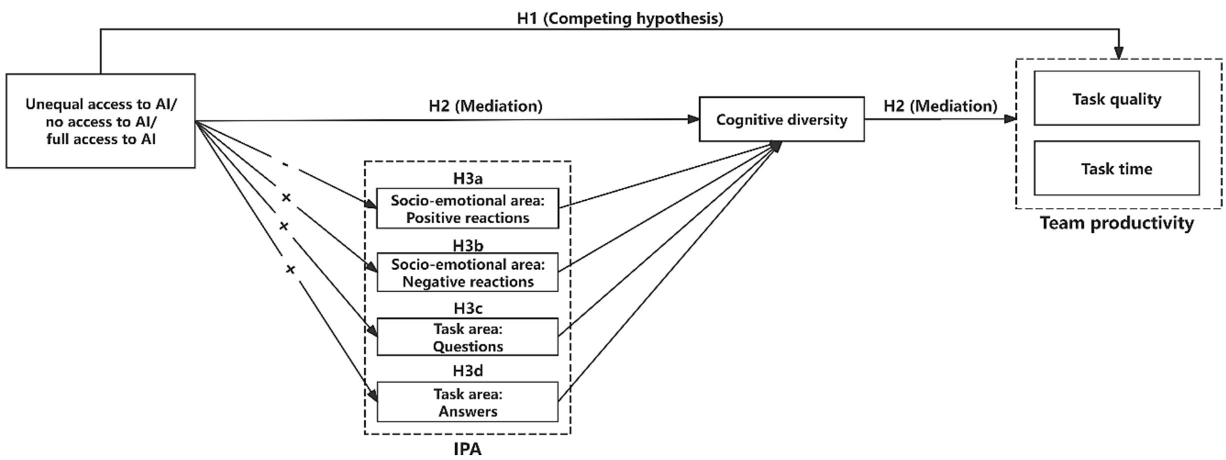


FIGURE 2
Research model.

students from various majors, who were randomly assigned to form 60 two-person teams. Each team first went through a control phase task where neither team member was permitted to use GenAI when completing the writing task (no access condition). Then, in the treatment phase, these teams were randomly assigned to one of two conditions: only one member could use GenAI (unequal access), or both members could use it (full access). For the GenAI tool, we employed Kimi 3.0, a Chinese-language-optimized large language model developed by Moonshot, chosen for its superior ability to handle lengthy text inputs and its suitability for the Chinese writing tasks.

3.1 Participants

This experiment involved 120 participants, who were made up of undergraduates and graduate students from a university in China. Our interest in studying face-to-face interactions in two-person teams made conducting the experiment with student samples the most feasible approach.

The sample of participants was 72% female, and a chi-square test of independence revealed that gender proportions did not significantly differ across the three experimental conditions, $\chi^2 = (2, N = 120) = 0.082$, $p = 0.960$. Approximately 18% of the sample were from the humanities and social sciences disciplines. Chi-square analysis indicated that the distribution of participants across the three conditions was statistically equivalent, $\chi^2 = (2, N = 120) = 0.000$, $p = 0.999$. Approximately 24% of the participants had prior experience related to marketing, with no significant differences observed across the three conditions, $\chi^2 = (2, N = 120) = 1.455$, $p = 0.483$. Approximately 95% of the participants had experience using generative AI tools (such as ChatGPT and Kimi), with no significant differences observed across the conditions, $\chi^2 = (2, N = 120) = 0.790$, $p = 0.674$. Participants received a reward of 50 RMB for participating in the experiment. Additionally, if their group's overall task quality was rated above 6 (on a range of 1–7, 7 the highest), each task would earn an extra 10 RMB.

3.2 Experiment procedure

This study selected Kimi 3.0, a large language model (LLM) developed by the Chinese company Moonshot, as the generative AI tool for team use for two advantages. First, Kimi was trained in and optimized for the Chinese language, making it an ideal choice given the designed writing task in Chinese. Second, Kimi outperforms other large models available in China in its ability to handle long texts (Team et al., 2025). This allows Kimi to better comprehend participants' extensive input commands and complete writing tasks more effectively.

Team activities were divided into five steps (shown in Figure 3): pre-test, control phase writing task, post-test 1, treatment phase writing task, and post-test 2. During the preparatory phase, participants completed an initial questionnaire to control for individual factors that could influence team communication and productivity, including demographic information, GenAI usage experience, and self-assessed skill levels in communication, creativity, and problem-solving. The first writing task served as the control task, where no members from any condition's teams could use GenAI. The second writing task served as the treatment task, where in condition 1, only one of two members was randomly assigned access to GenAI, and in condition 2, both individuals could use GenAI to complete the writing task. The first condition represented teams with unequal access to GenAI, while the second condition represented teams with full access to GenAI. Team members always have access to computers configured with task instruction documents and basic document editing tools. Only the individuals allowed to use GenAI were provided with a link to Kimi, and there was no restriction on how to interact with Kimi. All participants were not allowed to use any other websites or applications when not instructed to do so. After the experiment, we reviewed the on-site recordings to ensure that each group carried out the tasks in accordance with the above-mentioned requirements. The experiment design was approved by the university IRB (H20240616I).

3.3 Writing task design

The entire experiment comprises two writing tasks, in which two-person teams were asked to collaboratively produce a 700-character press release about a hypothetical product (an electric bicycle in the control phase and an AR glasses product in the treatment phase). This writing task is adapted from team collaboration tasks designed in prior literature (Noy and Zhang, 2023). Each writing task should not exceed 45 min in duration. Before starting each task, every team was first given basic information about the hypothesized product and writing instructions (see the Supplementary material Section 1 for Writing Tasks Instructions).

3.4 Measures

3.4.1 Access to GenAI

Access to GenAI serves as the main independent variable in the experiment. According to Hayes and Preacher (2014), we used indicator coding, also known as dummy coding, to represent this multi-categorical independent variable. To dummy-code three groups (no AI access, partial AI access, and full AI access), two dummy variables are constructed. The "No access" variable has a value of 1 if a case is in no access to the AI group and 0 otherwise. The "Full access" variable is set to 1 if a case is in the full AI access group and 0 otherwise. Partial AI access group functions as the reference category in the analysis and parameters reported in the model that are pertinent to group differences should be interpreted relative to this reference group.

3.4.2 Team productivity

Team productivity in this study was assessed along two key dimensions: task quality and task time, reflecting both the effectiveness and efficiency of team output (Harrison et al., 2003; Noy and Zhang, 2023).

3.4.2.1 Task quality

Following Noy and Zhang (2023), task quality was assessed by (blinded) expert raters working in marketing. Evaluators assigned an overall grade (1–7) to the writing task submissions based on three criteria: writing quality, content quality, and originality. Detailed

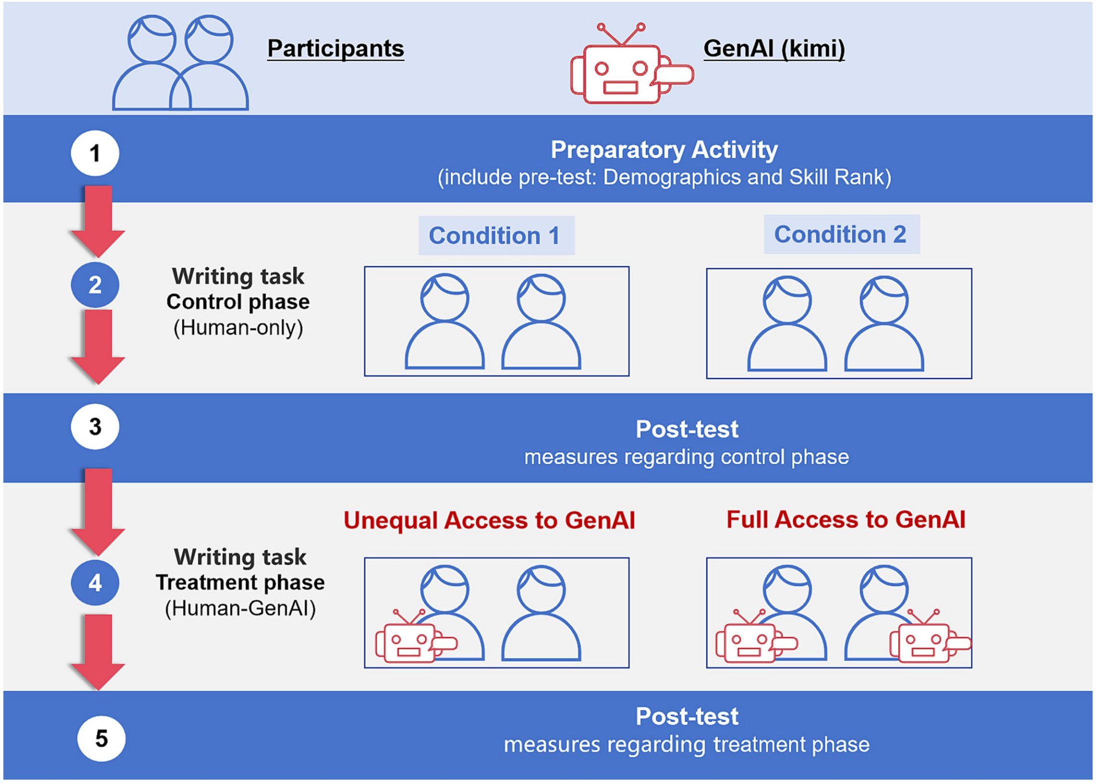


FIGURE 3
Experiment design.

instructions, including sample submissions with high and low scores, can be found in Section 2 in the Supplementary material. We recruited a total of nine professionals from the marketing industry as expert raters. Each of the 120 submissions was randomly assigned to three raters to ensure high reliability, with each rater evaluating 40 submissions. To encourage quality evaluations, raters were informed that their reward would be based on the correlation between their scores and those of the other raters. The Cronbach alpha between the three raters' scores was 0.791.

3.4.2.2 Task time

Task completion time was measured as the total duration each team spent working collaboratively on the assigned task. Following the procedures outlined by Noy and Zhang (2023), the entire task completion process was video-recorded for each team. Trained research assistants subsequently reviewed the recordings and extracted the task completion time for each team.

3.4.3 Cognitive diversity

To obtain an objective measure of cognitive diversity within each team, we employed a computational text analysis approach developed by Gurkan and Yan (2023). Team discussions were first transcribed from audio recordings, with manual corrections to ensure accuracy. We then identified and concatenated each team member's utterances across the entire team discussion. These text blocks were vectorized using the Universal Sentence Encoder (USE) from TensorFlow Hub (Cer et al., 2018), and each vector was normalized such that its magnitude (Euclidean norm) equals 1. Cognitive diversity was then calculated as the cosine distance (d) between the normalized vectors of the two team members (i and j), with higher values of d indicating greater cognitive dissimilarity. That is,

$$d\left(W_i, W_j\right) = 1 - \cos\left(W_i, W_j\right)$$

where W_i denotes the concatenated spoken text expressed by the individual i.

3.4.4 Team interaction process

Team interaction behaviors were coded using Bales' Interaction Process Analysis (IPA) framework, which includes 12 subcategories grouped into four functional areas: positive socio-emotional, negative socio-emotional, task-related answering, and task-related questioning (see Section 3 in the Supplementary material).

The unit of analysis was a single simple sentence or its equivalent—the smallest independent unit of meaning (Bales, 1950). Coders were instructed to treat short, complete responses (e.g., "Yes," "I agree") as standalone units. In contrast, sentence fragments that depend on preceding or following speech (e.g., "Because.," "And then.") should be merged with the adjacent utterance. Additionally, coders were trained to avoid combining sequential but distinct behaviors into a single code. For example, the utterance 'Yes, that makes sense, and what should we do next?' should be coded as two separate units—one for *Agreement* and one for *Asks for Suggestions*.

After the initial training, two coders independently coded 25% of the data. Discrepancies in this subset were discussed and resolved to refine the coding scheme for clarity. Once the coders achieved satisfactory agreement, they completed the remaining dataset. The final results yielded a Cohen's Kappa of 0.77, indicating substantial reliability. Any remaining disagreements were resolved through discussion, and consensus codes were used for analysis.

3.4.4.1 Socio-emotional area reactions

For each task group, the socio-emotional area behaviors were computed as relative frequency scores for the positive and negative interaction behaviors:

Positive socio-emotional behavior = Positive units/Total communication units.

Negative socio-emotional behavior = Negative units/Total communication units.

3.4.4.2 Task area reactions

For each task group, the task-related reactions were calculated as the ratio between two members' questioning or answering behaviors. Specifically, the ratio was determined by dividing the higher count of questioning behaviors by the lower count for each pair of team members. The formula used is as follows:

Concentrated questioning = max (Questioning units by A, Questioning units by B)/ min (Questioning units by A, Questioning units by B).

Concentrated answering = max (Answering units by A, Answering units by B)/ min (Answering units by A, Answering units by B).

Larger ratio scores indicate a higher level of concentration, meaning one member dominated that specific behavior (e.g., questioning or answering) to a greater extent. There exists a great imbalance between the two members in Q&A behaviors.

3.4.5 Control variables

We aggregated demographic variables to the team level, resulting in three control variables:

Female proportion. Proportion of female members in each team, calculated as the number of females divided by total team size (e.g., 0, 0.5, or 1 in two-person teams).

Marketing experience. If at least one member of a team has education or working experience in marketing-related education or work, this variable is marked as 1; if not, it is marked as 0.

Team skill. Participants were asked to rank their level in the following three teamwork skills: being an effective communicator, being creative and original, and problem-solving (Noy and Zhang, 2023). Each participant assigned a score of 3 to the skill they ranked first, 2 to the second, and 1 to the third. Based on these individual scores, we calculated team-level scores for each skill by averaging across team members, resulting in three variables: *team communication ability*, *team problem-solving ability*, and *team creativity*.

Due to concerns of multicollinearity (as the three scores are interdependent and sum to a constant), we included only *problem-solving skill* and *creativity* as control variables in our main analyses.

4 Results

To investigate how unequal access to AI predicts team task quality and completion time through team interaction processes and cognitive diversity, we conducted a PROCESS macro analysis. In this model, AI access (unequal access/ no access/ full access) served as the multi-categorical independent variable (IV); the four types of team interaction and cognitive diversity were included as mediators; and team productivity—task quality and task time—were treated as the dependent variables (DVs). We set the unequal AI access condition as the reference group and compared it with the no AI access condition (X_1) and the full AI access condition (X_2).

We first tested the hypothesized model, which demonstrated a good fit to the data: χ^2 (6, $N = 60$) = 4.243, $p = 0.644$. The probability that the root mean square error of approximation (RMSEA) is less than or equal to 0.05 was 0.900, and the other fit indices also indicated excellent model fit: comparative fit index (CFI) = 1.000, Tucker-Lewis index (TLI) = 1.027, and standardized root mean square residual (SRMR) = 0.0019. Having confirmed the overall model fit, we proceeded to examine each path in the model to evaluate our hypotheses.

4.1 Unequal access to GenAI leads to higher task quality and faster task completion

To examine the overall effect of AI integration structure on team productivity (H1), we compared task quality and completion time across the three AI integration conditions. As illustrated in Figure 4A, task quality in teams with unequal AI access ($M = 5.654$, $SD = 0.080$) was higher than in teams with no AI access ($M = 4.981$, $SD = 0.094$), $t = -5.103$, $p < 0.01$, and those with full AI access ($M = 5.250$, $SD = 0.118$), $t = 2.881$, $p < 0.01$. This finding is further supported by the OLS results reported in Model (1) of Table 1, where unequal AI access was associated with higher task quality compared to both the no access condition ($b = -0.673$, $p < 0.01$) and the full access condition ($b = -0.402$, $p < 0.05$).

A similar trend was observed in task completion time (Figure 4B). Teams with unequal AI access completed the task faster ($M = 10.013$, $SD = 0.807$) than human-only teams ($M = 30.483$, $SD = 1.108$; $t = 13.556$, $p < 0.01$), and those with full AI access ($M = 15.150$, $SD = 1.491$; $t = -3.312$, $p < 0.01$). These time savings are further reflected in the OLS estimates reported in Model (2) of Table 1, which show significantly reduced task duration for the unequal access condition compared to both no AI ($b = 20.752$, $p < 0.01$) and full AI access ($b = 5.982$, $p < 0.01$). Therefore, these findings provide support for H1a, indicating that unequal access to AI can significantly enhance team productivity by improving task quality and accelerating task completion.

4.2 Cognitive diversity links unequal AI access with enhanced task quality

To evaluate the hypothesized mediating role of cognitive diversity (H2), we first tested whether AI integration structure significantly influences cognitive diversity and whether cognitive diversity, in turn, predicts team productivity. Independent-sample t-tests were conducted to compare communication responses across three different AI access conditions, serving as a proxy for cognitive diversity. As shown in Figure 4C, teams with

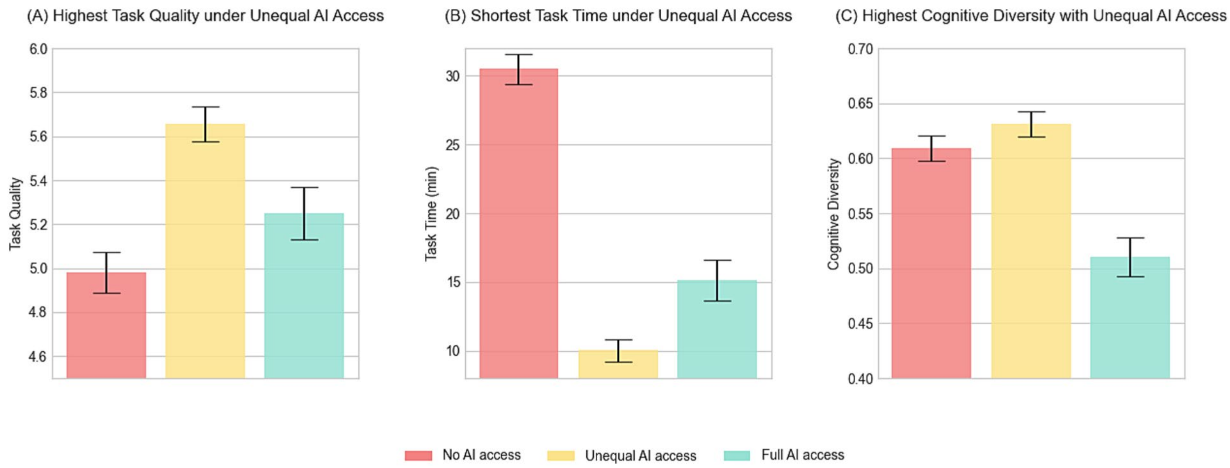


FIGURE 4
Team outcomes under different AI access conditions. Bar charts depict **(A)** task quality, **(B)** task completion time, and **(C)** cognitive diversity under three AI access conditions: no AI access, unequal AI access, and full AI access. Error bars reflect ±1 standard error of the mean.

TABLE 1 Regression of different AI access on team productivity.

Variable	(1) Task quality		(2) Task time	
	b	*SD*	*b*	*SD*
No access to AI	−0.6730***	−5.1990	20.7523***	14.1546
Full access to AI	−0.4023**	−2.2771	5.9820***	2.9896
Female proportion	−0.0123	−0.0593	1.3377	0.5705
Team problem-solving ability	0.0624	0.3415	−1.9904	−0.9619
Team creativity	0.2113	1.4935	−2.1255	−1.3264
Marketing experience	−0.0176	−0.1357	3.2350**	2.2036
Constant	5.1217***	10.1856	15.6672***	2.7509
N	120		120	
adj. R^2	0.170		0.644	

t statistics in parentheses. *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$.

unequal AI access exhibited significantly higher cognitive diversity ($M = 0.631$, $SD = 0.011$) than those with full AI access ($M = 0.510$, $SD = 0.018$), $t = 5.938$, $p < 0.01$. However, no significant difference was observed between the unequal AI access group and the no AI access group ($M = 0.609$, $SD = 0.012$), $t = -1.299$, $p = 0.197$. Further results from the mediation model (Figure 5) supported the pattern observed in the above findings. Compared to teams with unequal AI access (reference group), those with full AI access showed significantly lower cognitive diversity ($b = -0.107$, $SE = 0.025$, $p < 0.01$), while human-only teams did not differ significantly ($b = -0.017$, $SE = 0.020$, $p > 0.1$).

Importantly, the PROCESS model confirmed that cognitive diversity was positively associated with task quality ($b = 2.653$, $SE = 0.752$, $p < 0.01$), but showed no significant effect on task completion time ($b = 0.104$, $SE = 7.454$, $p > 0.1$). Bootstrapped indirect effect analysis (Table 2) further validated the mediating role of cognitive diversity: the indirect effect of unequal AI access (vs. full AI access) on task quality via cognitive diversity was significant ($b = -0.283$, $SE = 0.091$, 95% CI [−0.513, −0.134]). This suggests that unequal AI access can enhance team effectiveness by fostering greater cognitive diversity. In sum, these findings support H2 by demonstrating that cognitive diversity significantly mediates the relationship between AI integration structure and team productivity—specifically, by enhancing task quality.

4.3 Mechanisms underlying the impact of GenAI access on cognitive diversity and team productivity

4.3.1 Socio-emotional area: negative reactions and cognitive diversity act as serial mediators between unequal AI access and task quality

H3a and H3b proposed that socio-emotional team interactions—positive and negative reactions—mediate the relationship between AI integration structure and cognitive diversity. Path analyses revealed that unequal AI access significantly increased negative socio-emotional behaviors compared to full AI access ($b = -0.028$, $SE = 0.007$, $p < 0.01$), which, in turn, positively influenced cognitive diversity ($b = 0.811$, $SE = 0.242$, $p < 0.01$), supporting H3b. However, no significant effects were found for AI integration structure on positive reactions (relative to no access: $b = 0.023$, $SE = 0.014$, $p > 0.05$; relative to full access: $b = -0.005$, $SE = 0.017$, $p > 0.05$), thus failing to support H3a.

As previously demonstrated, cognitive diversity mediates the relationship between unequal AI access and task quality. Building on this, we further tested whether negative socio-emotional interactions contribute to this indirect pathway. Results from the PROCESS model (Table 2) showed a significant bootstrapped serial indirect effect involving AI integration, negative socio-emotional behaviors, cognitive diversity, and task quality (relative to full AI access: $b = -0.061$, $SE = 0.029$, 95% CI [−0.148, −0.021]). These findings suggest that unequal AI access can enhance task quality by increasing negative socio-emotional reactions, which in turn promote greater cognitive diversity. In

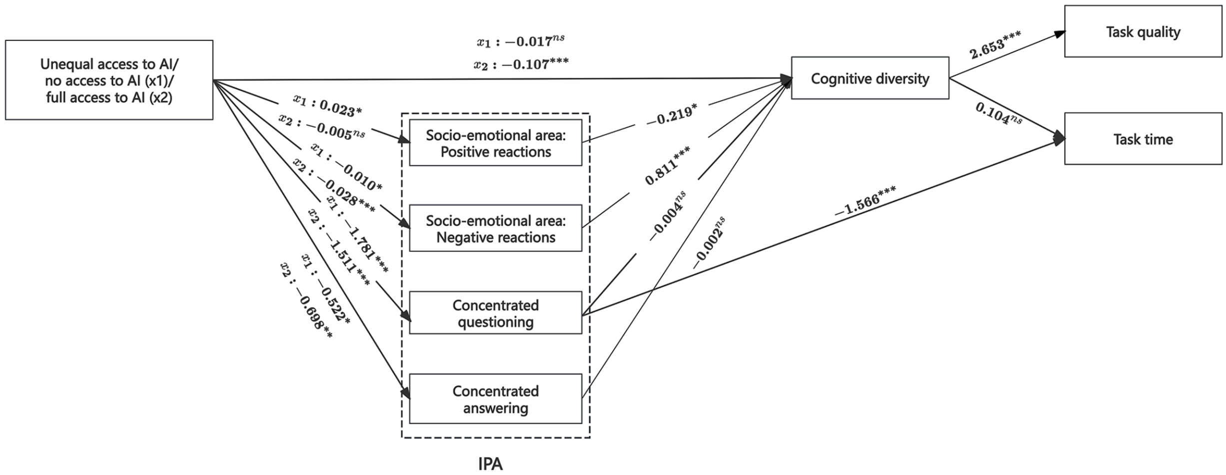


FIGURE 5
Results of the mediation model (with unequal access condition as reference group). ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$, ns, not significant. Other non-significant paths are omitted in the figure.

TABLE 2 Significant indirect effects tested by 95% bootstrapped confidence intervals.

Indirect path	Estimated effect	95%CI
Unequal AI access (versus full AI access) → cognitive diversity → task quality	−0.283	[−0.513, −0.134]
Unequal AI access (versus full AI access) → negative socio-emotional reactions → cognitive diversity → task quality	−0.061	[−0.148, −0.021]
Unequal AI access (versus no AI access) → concentrated task-related questioning → task time	2.790	[1.628, 4.138]
Unequal AI access (versus full AI access) → concentrated task-related questioning → task time	2.367	[0.780, 4.578]

other words, negative interpersonal communication and cognitive diversity function as sequential mediators linking unequal AI access to improved team productivity.

4.3.2 Task area: concentrated questioning mediates the relationship between unequal AI access and task time

To test the role of task-related interactions, H4a and H4b focused on whether concentrated questioning and answering mediate the link between AI integration and cognitive diversity. Path analysis (Figure 5) indicated that unequal AI access significantly increased both concentrated questioning (relative to no access: $b = -1.781$, $SE = 0.381$, $p < 0.01$; relative to full access: $b = -1.511$, $SE = 0.452$, $p < 0.01$) and concentrated answering behaviors (relative to no access: $b = -0.522$, $SE = 0.278$, $p < 0.1$; relative to full access: $b = -0.698$, $SE = 0.295$, $p < 0.05$). While the direction of these relationships aligned with our assumptions, the mediating effects did not. Concentrated task-related questioning ($b = -0.004$, $SE = 0.005$, $p > 0.1$) and answering behavior ($b = -0.002$, $SE = 0.007$, $p > 0.1$) showed no significant impact on cognitive diversity, contrary to H3c and H3d.

Although task-related behaviors did not mediate the relationship between AI integration and cognitive diversity, we found that concentrated questioning had a direct negative effect on task time ($b = -1.566$, $SE = 0.502$, $p < 0.01$). The bootstrapped indirect effects (Table 2) from unequal AI access to task time through concentrated questioning were significant (relative to no access: $b = 2.790$, $SE = 0.658$, 95% CI [1.628, 4.138]; relative to full access: $b = 2.367$, $SE = 0.702$, 95% CI [0.780, 4.578]). These findings suggest that unequal AI access can shorten task time by prompting non-AI users to take on a greater share of task-related questioning, thereby increasing the efficiency of team interactions.

4.4 Robustness check

This study presents two additional analyses to strengthen the robustness of the findings reported above.

4.4.1 Baseline team characteristics comparison

To ensure that there are no significant differences in team baseline characteristics across conditions and to rule out the impact of initial levels on the observed outcomes, t-tests were conducted on various team characteristics. Table 3 provides descriptive statistics for team-level control variables, task quality, task time, cognitive diversity, and four categories of team interaction processes during the control phase. t tests compared these variables across two conditions and found no significant differences in terms of baseline team characteristics.

4.4.2 Measuring task quality by originality

In the main analysis, task quality was assessed through three dimensions—content quality, writing quality, and originality (see section 3.4.2). Given that cognitive diversity is widely acknowledged to influence team creativity (Mathuki and Zhang, 2024; Qi et al., 2022; Wang et al., 2016), we identified originality as the core dimension most directly driven by cognitive diversity. To test the robustness of our findings, we re-ran the mediation analysis using originality as the sole indicator of task quality.

The model exhibited a good fit (χ^2 (6, N = 60) = 2.243, p = 0.644). As shown in Figure 6, cognitive diversity had a significant positive effect on originality (b = 4.619, SE = 1.346, p < 0.01). Moreover, the serial mediation pathway from unequal AI access to originality—via negative socio-emotional interactions and cognitive diversity—was also significant, relative to full AI access (b = −0.105, SE = 0.050, 95% CI [−0.248, −0.036]). These results provide robust support for our earlier conclusions, reaffirming that unequal access to GenAI enhances task quality primarily through its effect on team interaction dynamics and cognitive diversity, particularly as reflected in originality.

4.4.3 Measuring concentrated task-related behavior using difference scores

To test whether our findings are sensitive to how Concentrated Task-Related Behavior is measured, we re-estimated the model using an alternative operationalization based on difference scores (De Jong et al., 2022).

Concentrated questioning = |Questioning units by A − Questioning units by B| / Total questioning units.

Concentrated answering = |Answering units by A − Answering units by B| / Total Answering units.

A and B represent the two team members. Higher values indicate a greater concentration of the corresponding behavior within teams.

The results (Figure 7) show that the overall model fit remained acceptable under this alternative specification (χ^2 (6, N = 60) = 5.752, p = 0.452). Importantly, the hypothesized indirect path from unequal AI access to task time via concentrated questioning behavior remained statistically significant (relative to no access: b = 2.913, SE = 1.138, 95% CI [1.003, 5.399]; relative to full access: b = 2.390, SE = 1.059, 95% CI [0.778, 4.939]), supporting the robustness of the proposed mechanism.

5 Discussion

5.1 Key findings

Our analysis revealed four key patterns. First, contrary to a general intuition that fully equipping working teams with GenAI could enhance team productivity, we observe that teams with unequal AI access actually improved task quality by improving team cognitive diversity. Though unequal AI access does not seem to affect task time. Second, when examining team interactions, unequal AI access also had some interesting effects. In the socio-emotional area interactions, it sparked more negative reactions, like disagreement, but did not really change how often people expressed positive emotions. In the task area interactions, it led to more concentrated task-related questioning and answering, with certain team members taking the lead in asking questions and others concentrating on answering them. Third, more concentrated task-related questioning explains why unequal AI access (versus full and no access) reduced task time. That is, when a subset of team members primarily handles questioning, task completion accelerates. Fourth, there exists a positive serial mediation path from unequal AI access (versus full access) to improved task quality, sequentially through increased negative socio-emotional behaviors and greater cognitive diversity. In other words, although unequal access led to more disagreement among team members, this also encouraged a broader range of thinking styles—ultimately helping the team perform better.

5.2 Theoretical implications

This study offers three key theoretical contributions to the literature on AI integration structures and team processes in HATs. First, we clarify conflicting perspectives on the relationship between unequal AI access and team productivity through the lens of cognitive diversity. We find that cognitive diversity induced by partial AI access enhances task quality, aligning with previous

TABLE 3 Descriptive statistics.

Variable	Condition 1 (unequal access) (N = 40) Mean (SD)	Condition 2 (Full access) (N = 20) Mean (SD)	t tests t (p)
Female proportion	0.725 (0.054)	0.700 (0.076)	0.269 (0.789)
Marketing experience	0.400 (0.496)	0.200 (0.410)	1.555 (0.126)
Team problem-solving ability	1.763 (0.059)	1.925 (0.098)	−1.495 (0.140)
Team creativity	2.075 (0.085)	2.000 (0.115)	0.517 (0.607)
Task quality	4.958 (0.818)	5.025 (0.508)	−0.333 (0.740)
Task time	30.550 (9.419)	30.350 (6.831)	0.084 (0.933)
Cognitive diversity	0.618 (0.102)	0.591 (0.063)	1.103 (0.274)
Positive socio-emotional reactions	0.142 (0.058)	0.155 (0.060)	−0.866 (0.390)
Negative socio-emotional reactions	0.027 (0.028)	0.025 (0.024)	0.268 (0.790)
Concentrated task-related questioning	1.347 (0.056)	1.297 (0.055)	0.560 (0.578)
Concentrated task-related answering	2.186 (0.205)	1.781 (0.230)	1.217 (0.229)

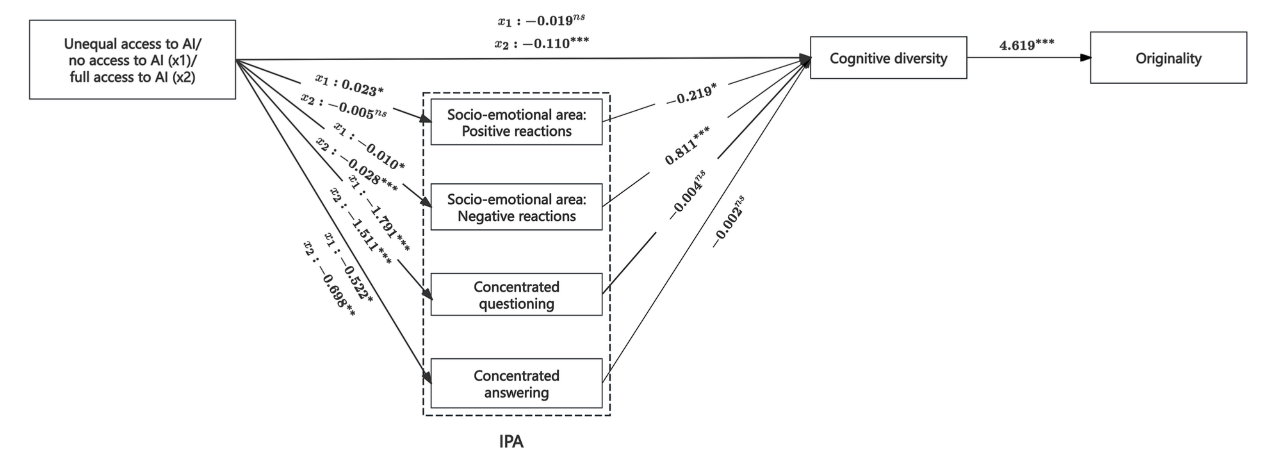


FIGURE 6
Results of the mediation model (with task quality measured by originality). ***p < 0.01, **p < 0.05, *p < 0.1, ns, not significant. Paths related to task time are omitted for clarity.

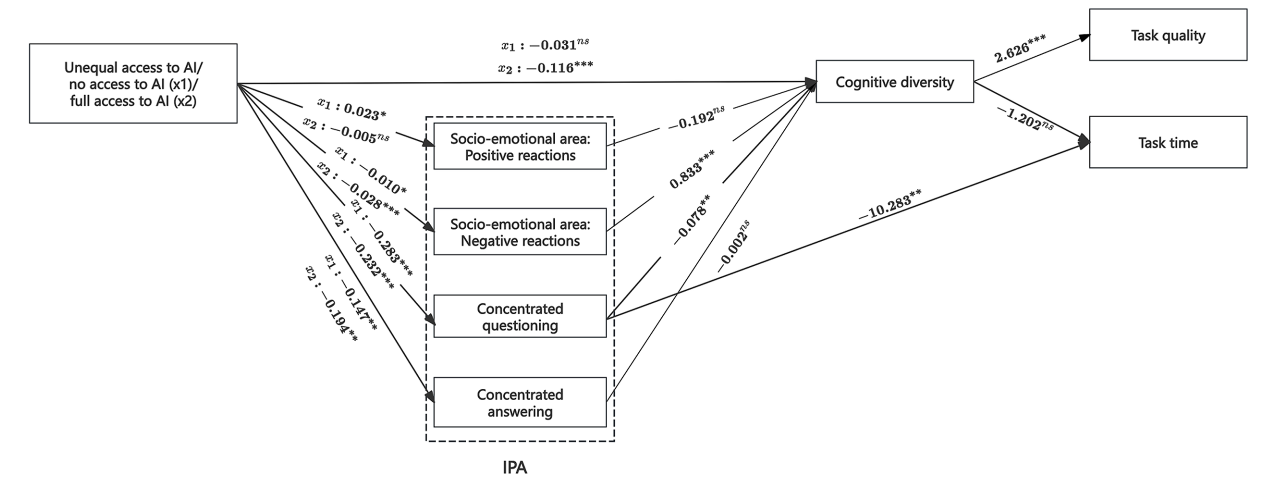


FIGURE 7
Results of the mediation model (with concentrated task-related behavior measured using difference scores). ***p < 0.01, **p < 0.05, *p < 0.1, ns, not significant. Other non-significant paths are omitted in the figure.

findings by Wang et al. (2016) and Aggarwal and Woolley (2019), which emphasize the value of diverse perspectives in team collaboration. In terms of task completion time, cognitive diversity has no significant impact, in contrast to prior literature that documented both its positive (Li et al., 2022; Pieterse et al., 2011) and negative (Mohammed and Schillinger, 2022; Narayan et al., 2021) effect on team working efficiency. Thus, we emphasize the role of cognitive diversity as a key mediator through which unequal AI access improves the quality of creative task outputs.

Second, by building and testing an I-P-S-O model, we theorize and empirically demonstrate that unequal AI access gives rise to distinctive interaction processes (P factor) and emergent cognitive states (S factor), which sequentially mediate its impact on teaming effectiveness. This contributes to team science literature by identifying the underlying mechanisms through which inconsistent technological usage shapes collaborative dynamics. Moving beyond the view of AI as a uniform group-level resource (e.g., Gurkan and Yan, 2023), we demonstrate how individual-level differences in AI technology access may actively reshape information distribution within teams. Specifically, we find that unequal AI access alters the flow of communication by concentrating questioning and answering behaviors within certain members. In other words, when AI access is unequal, information flows become more fixed: some members possess more task-relevant information and thus predominantly answer questions, while others, lacking such information, primarily ask questions. This pattern corresponds to the I-P path of our I-P-S-O model. Furthermore, informational asymmetry caused by unequal AI integration fosters deeper discussions, thereby enhancing team cognitive diversity. This reflects the I-S path in our model. Our findings show that unequal access stimulates more diverse perspectives, whereas full access may have a homogenizing effect by leading team members to base their reasoning and decisions on similar AI-generated inputs. This pattern aligns with prior research on social confirmation bias (Lu et al., 2012; Stasser and Titus, 2003), which shows that shared information among members can overshadow unique contributions and suppress cognitive diversity.

Moreover, this study finds that cognitive diversity can emerge dynamically from team interaction processes, which refers to the P-S link in the IPSO model. We empirically identify negative socio-emotional behaviors, especially disagreement, in team communication that are most strongly associated with the emergence of cognitive diversity. This finding supports theoretical propositions by Marks et al. (2001) and Mello and Rentsch (2015), who suggested that cognitive diversity functions as an emergent property shaped by ongoing team dynamics. Our study thus offers empirical insight into the interpersonal communicative mechanisms underpinning the development of cognitive diversity in human-AI teams.

Third, our study also explores the direct effects of varied aspects of team interaction dynamics on team productivity. Unlike prior studies that treat interactions as a general concept (Gurkan and Yan, 2023; Mello and Rentsch, 2015), we differentiate interactions in the task and socio-emotional domains and find that they each have distinct effects on task completion time and task quality, respectively. In the task domain, team interactions characterized by concentrated patterns of questioning are closely associated with shorter task completion times. These patterns only arise when the GenAI access is partial, meaning that AI users may tend to provide orientation and information, while non-users seek suggestions and ask more questions. Interestingly, only concentrated task-related questioning—rather than answering—appears to accelerate task completion. This may be explained by De Jong et al. (2022), who argue that questioning can signal recognition of others' expertise or leadership, suggesting that AI access may function as a status characteristic, reinforcing status hierarchies and improving decision-making efficiency. Both theoretical explanations offer interesting insights worthy of future empirical testing.

In the socio-emotional domain, negative interactions—particularly those stemming from disagreement under conditions of unequal AI access—are found to have a positive impact on task quality. While this finding partially aligns with prior research (Mesmer-Magnus and DeChurch, 2009; Stasser and Titus, 1985; Van Knippenberg and Schippers, 2007), which highlights that uneven information distribution can lead to conflict, those studies typically view such conflict as detrimental to team cohesion and performance. In contrast, our findings suggest that task-related disagreement, though seemingly negative, may stimulate deeper cognitive engagement and enhance team outcomes. This supports the view of Farh et al. (2010), who argue that moderate task conflict can benefit collaboration and creativity.

In conclusion, our IPSO model proposes a comprehensive influence pathway—from AI integration structure as a team input, through observable team interaction behaviors and cognitive emergent states, to team productivity such as task quality and completion time. This enriches the IMO model for HATs proposed by O'Neill et al. (2023), providing a theoretically grounded explanation of how varying levels of GenAI access shape emergent cognition and collaborative performance. Our findings offer a foundation for developing strategic GenAI integration frameworks to optimize human-agent collaboration in diverse team environments.

5.3 Practical implications

In practical terms, this paper provides assistance and guidance for establishing management strategies for short-term Human-GenAI teams. The two-person teams in this study can be expanded to multi-person teams in the real world. We demonstrate that unequal GenAI access among team members can reshape information flows and influence team cognitive diversity, thereby impacting task quality. Rather than simply pursuing equal access across all members, organizations should consider the strategic allocation of GenAI based on task requirements, member roles, and the desired level of cognitive diversity. For instance, in short-term collaborative tasks that require innovative problem-solving—such as brainstorming sessions or team debates—a certain level of cognitive divergence resulting from differentiated AI usage may be beneficial. However, for teams that emphasize long-term relationships and the personal development of members, alternative allocation strategies may be more appropriate. By strategically limiting access to GenAI, organizations can potentially harness the strengths of both human expertise and AI capabilities, fostering an environment in which diverse perspectives contribute to both task outcomes and team development.

In light of our findings, team leaders and facilitators should actively monitor and manage interaction patterns that emerge from unequal GenAI integration. Our results suggest that disagreement stemming from unequal AI distribution is not inherently detrimental; in fact, it significantly enhances team cognitive diversity, which in turn improves the quality of team output. Therefore, when task-related disagreements arise between AI users and non-users, managers need not suppress such conflict. Instead, they should view it as a potential catalyst for creativity, intervening only to guide it constructively. However, when such task conflict escalates into relationship conflict or fosters mistrust among members, targeted interventions become necessary to maintain psychological safety and team cohesion.

5.4 Limitations and future directions

While our study provides valuable insights into team cognition states under varied AI integration structures, several limitations should be acknowledged to inform future research and deepen understanding of the topic. First, our sample was somewhat limited in its diversity due to practical constraints in recruiting participants for a controlled laboratory experiment involving face-to-face team interactions. Recruiting student participants was the most feasible and appropriate approach given resource availability and the need for experimental control. While the student sample included individuals from a broad range of academic disciplines, reflecting some diversity in cognitive and educational backgrounds, it is important to note that these participants generally lack substantial real-world work experience and exposure to professional team environments. This limitation may affect the external validity of our findings.

Second, there was a notable gender imbalance in our sample. Prior research suggests gender can influence perceptions of status (Levin, 2004; Ridgeway, 2001), communication style (Furumo and Pearson, 2007), and participation equity within teams (Bear and Woolley, 2011). Although we conducted additional post-hoc analyses and found no significant gender differences in team processes or productivity across different AI access conditions, this issue warrants further investigation. Future research should seek more gender-balanced samples to ensure robustness and broader applicability of the findings.

Third, the design of the team task has certain constraints. In this study, participants were required to complete the task through real-time communication within a limited amount of time, a format that mirrors many real-world settings, such as problem-solving meetings or short-term team competitions. However, the impact of AI integration structures in long-term collaboration remains an important area of exploration. In extended projects, task roles tend to be more clearly defined, and learning processes become more prominent. Emotional connections among team members may also deepen. Whether unequal AI access continues to outperform full AI access in fostering cognitive development in such contexts is a question worth investigating. Additionally, the control and treatment tasks used different product prompts—an electric bicycle and AR glasses, respectively. Although both prompts were pre-tested by domain experts to ensure similar levels of difficulty, complexity, and creative demand, this variation may still introduce uncontrolled differences in team performance. This design decision aimed to reduce learning and fatigue effects from task repetition, but future research would benefit from employing counterbalanced or equivalent task designs to further validate the robustness of the findings.

Finally, our study used a text-to-text interaction modality when prompting AI. Although this is currently the most mainstream interaction modality, future team collaboration may involve multimodal interactions, such as voice-based communication. It remains an open question whether multimodal interfaces could reduce the asymmetry in information and perceived status brought about by unequal AI access, thereby influencing interaction behaviors and cognitive states differently. This presents a promising direction for future research.

6 Conclusion

Team cognitive emergent states have long been recognized as critical components of team processes. This study explores how varied GenAI integration structures within HATs influence team cognitive diversity and, in turn, affect team productivity in areas such as task quality and efficiency. By uncovering the behavioral mechanisms—such as disagreement—that link AI access to divergent communication, this research deepens the understanding of how cognitive diversity emerges under unequal AI access. These differences in team cognition significantly enhance team output quality. Overall, this study highlights the central role of interaction dynamics and cognitive diversity in shaping team outcomes under varying patterns of GenAI use. Future work should continue to examine the nuanced mechanisms and interaction mode behind GenAI integration to better support collaborative performance in increasingly hybrid human-AI environments.

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 humans were approved by Shanghai Jiao Tong University Institutional Review Board for Human Research Protections. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

JH: Methodology, Data curation, Investigation, Conceptualization, Visualization, Writing – original draft, Formal analysis. RR: Writing – review & editing, Conceptualization, Funding acquisition, Supervision, Methodology, Project administration.

Funding

The author(s) declare that no financial support was received for the research and/or publication of this article.

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.

Generative AI statement

The authors declare that no Gen AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2025.1636906/full#supplementary-material

References

Aggarwal, I., and Woolley, A. W. (2019). Team creativity, cognition, and cognitive style diversity. *Manag. Sci.* 65, 1586–1599. doi: 10.1287/mnsc.2017.3001

Al Naqbi, H., Bahroun, Z., and Ahmed, V. (2024). Enhancing work productivity through generative artificial intelligence: a comprehensive literature review. *Sustainability* 16:1166. doi: 10.3390/su16031166

Bales, R. F. (1950). Interaction process analysis; a method for the study of small groups. Cambridge, MA: Addison-Wesley.

Bankins, S., and Formosa, P. (2023). The ethical implications of artificial intelligence (AI) for meaningful work. *J. Bus. Ethics* 185, 725–740. doi: 10.1007/s10551-023-05339-7

Bayerl, P. S., Lauche, K., and Axtell, C. (2016). Revisiting group-based technology adoption as a dynamic process. *MIS Q.* 40, 775–784. doi: 10.25300/MISQ/2016/40.3.12

Bear, J. B., and Woolley, A. W. (2011). The role of gender in team collaboration and performance. *Interdiscip. Sci. Rev.* 36, 146–153. doi: 10.1179/030801811X13013181961473

Berger, J., Rosenholtz, S. J., and Zelditch, M. (1980). Status organizing processes. *Annu. Rev. Sociol.* 6, 479–508. doi: 10.1146/annurev.so.06.080180.002403

Becker, A., Carte, T., and Chidambaram, L. (2008). "The effects of collaborative technology appropriation on group outcomes", in *DIGIT 2008 Proceedings*

Bunderson, J. S., and Reagans, R. E. (2011). Power, status, and learning in organizations. *Organ. Sci.* 22, 1182–1194. doi: 10.1287/orsc.1100.0590

Bunderson, J. S., and Sutcliffe, K. M. (2002). Comparing alternative conceptualizations of functional diversity in management teams: process and performance effects. *Acad. Manag. J.* 45, 875–893. doi: 10.2307/3069319

BusinessWire. (2023). AI at work: It's Here and it's working, whether you know it or not. Available online at: https://www.businesswire.com/news/home/20231024605486/en/AI-at-Work-Its-Here-and-Its-Working-Whether-You-Know-It-or-Not (Accessed April 17).

Cer, D., Yang, Y., Kong, S.-y., Hua, N., Limtiaco, N., John, R. S., et al. (2018). Universal sentence encoder. *arXiv:1803.11175*. doi: 10.48550/arXiv.1803.11175

Chung, C., and Pennebaker, J. (2011). "The psychological functions of function words" in Social communication. (New York, NY: Psychology Press), 343–359.

Correll, S. J., and Ridgeway, C. L. (2003). "Expectation states theory" in Handbook of social psychology. (New York, NY: Springer), 29–51.

Cronin, M. A., Weingart, L. R., and Todorova, G. (2011). Dynamics in groups: are we there yet? *Acad. Manage. Ann.* 5, 571–612. doi: 10.5465/19416520.2011.590297

Cui, K. Z., Demirer, M., Jaffe, S., Musolff, L., Peng, S., and Salz, T. (2024). "The Productivity Effects of Generative AI: Evidence from a Field Experiment with GitHub Copilot", in: An MIT Exploration of Generative AI. doi: 10.21428/e4baedd9.3ad85f1c

De Jong, W., Brouwer, J., Lockhorst, D., De Kleijn, R., Van Tartwijk, J., and Noordegraaf, M. (2022). Describing and measuring leadership within school teams by applying a social network perspective. *Int. J. Educ. Res. Open* 3:100116. doi: 10.1016/j.ijedro.2021.100116

DeSanctis, G., and Poole, M. S. (1994). Capturing the complexity in advanced technology use: adaptive structuration theory. *Organ. Sci.* 5, 121–147. doi: 10.1287/orsc.5.2.121

Doshi, A. R., and Hauser, O. P. (2024). Generative AI enhances individual creativity but reduces the collective diversity of novel content. *Sci. Adv.* 10:eadn5290. doi: 10.1126/sciadv.adn5290

Farh, J.-L., Lee, C., and Farh, C. I. (2010). Task conflict and team creativity: a question of how much and when. *J. Appl. Psychol.* 95, 1173–1180. doi: 10.1037/a0020015

Finholt, T., Sproull, L., and Kiesler, S. (2014). "Communication and performance in ad hoc task groups", in Intellectual Teamwork. (New York, NY: Psychology Press), 291–325.

Furumo, K., and Pearson, J. M. (2007). Gender-based communication styles, trust, and satisfaction in virtual teams. *J. Informat. Informat. Technol. Organiz. (Years 1–3)* 2, 047–060. doi: 10.28945/138

Gohar, K., and Utley, K.. (2023). Don't let gen AI limit your Team's creativity. Available online at: https://hbr.org/2024/03/dont-let-gen-ai-limit-your-teams-creativity (Accessed May 26).

Gomez, C. J., and Lazer, D. M. (2019). Clustering knowledge and dispersing abilities enhances collective problem solving in a network. *Nat. Commun.* 10:5146. doi: 10.1038/s41467-019-12650-3

Gurkan, N., and Yan, B. (2023). Chatbot catalysts: Improving Team decision-making through cognitive diversity and information elaboration. Proceedings of the 44th International Conference on Information Systems (ICIS 2023). Atlanta, GA: Association for Information Systems.

Hackman, J. R. (1978). The design of work in the 1980s. *Organ. Dyn.* 7, 3–17. doi: 10.1016/0090-2616(78)90031-1

Han, Y., Qiu, Z., Cheng, J., and Lc, R.. (2024). When teams embrace AI: human collaboration strategies in generative prompting in a creative design task. Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems.

Harrison, D. A., and Klein, K. J. (2007). What's the difference? Diversity constructs as separation, variety, or disparity in organizations. *Acad. Manag. Rev.* 32, 1199–1228. doi: 10.5465/amr.2007.26586096

Harrison, D. A., Mohammed, S., McGrath, J. E., Florey, A. T., and Vanderstoep, S. W. (2003). Time matters in team performance: effects of member familiarity, entrainment, and task discontinuity on speed and quality. *Pers. Psychol.* 56, 633–669. doi: 10.1111/j.1744-6570.2003.tb00753.x

Harrison, D. A., Price, K. H., Gavin, J. H., and Florey, A. T. (2000). Time, teams, and task performance: a longitudinal study of the changing effects of diversity on group functioning. *Acad. Manage. Proc.* 2000, C1–C6. doi: 10.5465/apbpp.2000.5439172

Hayes, A. F., and Preacher, K. J. (2014). Statistical mediation analysis with a multicategorical independent variable. *Br. J. Math. Stat. Psychol.* 67, 451–470. doi: 10.1111/bmsp.12028

Homan, A. C., Van Knippenberg, D., Van Kleef, G. A., and De Dreu, C. K. (2007). Interacting dimensions of diversity: cross-categorization and the functioning of diverse work groups. *Group Dyn. Theory Res. Pract.* 11:79. doi: 10.1037/1089-2699.11.2.79

Horwitz, S. K., and Horwitz, I. B. (2007). The effects of team diversity on team outcomes: a meta-analytic review of team demography. *J. Manag.* 33, 987–1015. doi: 10.1177/0149206307308587

Humlum, A., and Vestergaard, E. (2025). The unequal adoption of ChatGPT exacerbates existing inequalities among workers. *Proc. Natl. Acad. Sci.* 122:e2414972121. doi: 10.1073/pnas.2414972121

Ilgen, D. R., Hollenbeck, J. R., Johnson, M., and Jundt, D. (2005). Teams in organizations: from input-process-output models to IMOI models. *Annu. Rev. Psychol.* 56, 517–543. doi: 10.1146/annurev.psych.56.091103.070250

Isohätälä, J., Näykki, P., and Järvelä, S. (2020). Cognitive and socio-emotional interaction in collaborative learning: exploring fluctuations in students' participation. *Scand. J. Educ. Res.* 64, 831–851. doi: 10.1080/00313831.2019.1623310

Jeong, J., and Jeong, I. (2024). Driving creativity in the AI-enhanced workplace: roles of self-efficacy and transformational leadership. *Curr. Psychol.* 44, 8001–8014. doi: 10.1007/s12144-024-07135-6

Joniaková, Z., Jankelová, N., Blštáková, J., and Némethová, I. (2021). Cognitive diversity as the quality of leadership in crisis: Team performance in health service during the COVID-19 pandemic. *Healthcare* 9:313. doi: 10.3390/healthcare9030313

Kennedy, K. A., and Pronin, E. (2008). When disagreement gets ugly: perceptions of bias and the escalation of conflict. *Personal. Soc. Psychol. Bull.* 34, 833–848. doi: 10.1177/0146167208315158

Korzynski, P., Mazurek, G., Altmann, A., Ejdys, J., Kazlauskaite, R., Paliszkiewicz, J., et al. (2023). Generative artificial intelligence as a new context for management theories: analysis of ChatGPT. *Cent. Eur. Manag. J.* 31, 3–13. doi: 10.1108/CEMJ-02-2023-0091

Kozlowski, S. W., and Ilgen, D. R. (2006). Enhancing the effectiveness of work groups and teams. *Psychol. Sci. Public Interest* 7, 77–124. doi: 10.1111/j.1529-1006.2006.00030.x

Kurtzberg, T. R. (2005). Feeling creative, being creative: an empirical study of diversity and creativity in teams. *Creat. Res. J.* 17, 51–65. doi: 10.1207/s15326934crj1701_5

Kwarteng, S., Frimpong, S. O., Asare, R., and Wiredu, T. J. N. (2023). Effect of employee recognition, employee engagement on their productivity: the role of transformational leadership style at Ghana health service. *Curr. Psychol.* 43, 5502–5513. doi: 10.1007/s12144-023-04708-9

Levin, S. (2004). Perceived group status differences and the effects of gender, ethnicity, and religion on social dominance orientation. *Polit. Psychol.* 25, 31–48. doi: 10.1111/j.1467-9221.2004.00355.x

Li, Y., Feng, H., Peng, Z., Zhou, L., and Wan, J. (2022). Diversity-aware unmanned vehicle team arrangement in mobile crowdsourcing. *EURASIP J. Wirel. Commun. Netw.* 2022:56. doi: 10.1186/s13638-022-02139-x

Li, N., Zhou, H., and Mikel-Hong, K. (2024). Generative AI enhances team performance and reduces need for traditional teams. *arXiv:2405.17924*. doi: 10.48550/arXiv.2405.17924

Liu, J. K., Wei, V. K., and Wong, D. S. (2004). Linkable spontaneous anonymous group signature for ad hoc groups. Australasian Conference on Information Security and Privacy,

Lu, L., Yuan, Y. C., and McLeod, P. L. (2012). Twenty-five years of hidden profiles in group decision making: a meta-analysis. *Personal. Soc. Psychol. Rev.* 16, 54–75. doi: 10.1177/1088868311417243

Majchrzak, A., More, P. H., and Johnson, J. (2012) Practices to improve group creativity: a longitudinal field investigation. 2012 45th Hawaii International Conference on System Sciences.

Mannes, A. E., Soll, J. B., and Larrick, R. P. (2014). The wisdom of select crowds. *J. Pers. Soc. Psychol.* 107, 276–299. doi: 10.1037/a0036677

Marks, M. A., Mathieu, J. E., and Zaccaro, S. J. (2001). A temporally based framework and taxonomy of team processes. *Acad. Manag. Rev.* 26, 356–376. doi: 10.2307/259182

Mathuki, E., and Zhang, J. (2024). Cognitive diversity, creativity and team effectiveness: the mediations of inclusion and knowledge sharing. *Vine J. Inf. Knowl. Manag. Syst.* 54, 1379–1403. doi: 10.1108/VJIKMS-06-2022-0190

McNeese, N. J., Demir, M., Cooke, N. J., and Myers, C. (2018). Teaming with a synthetic teammate: insights into human-autonomy teaming. *Hum. Factors* 60, 262–273. doi: 10.1177/0018720817743223

Meeussen, L., and Van Dijk, H. (2016). The perceived value of team players: a longitudinal study of how group identification affects status in work groups. *Eur. J. Work Organ. Psychol.* 25, 200–211. doi: 10.1080/1359432X.2015.1028378

Mello, A. L., and Rentsch, J. R. (2015). Cognitive diversity in teams. *Small Group Res.* 46, 623–658. doi: 10.1177/1046496415602558

Mesmer-Magnus, J. R., and DeChurch, L. A. (2009). Information sharing and team performance: a meta-analysis. *J. Appl. Psychol.* 94, 535–546. doi: 10.1037/a0013773

Miller, C., Burke, L., and Glick, W. C. (1998). Cognitive diversity among upper-echelon executives: implications for strategic decision processes. *Strateg. Manage. J.* 19, 39–58.

Mohammed, S., Hamilton, K., Marhefka, J., Tirrell, B., Davis, C., and Hong, H. (2023). To share or not to share? Knowledge convergence and divergence in cross-disciplinary collaboration. *J. Organ. Psychol.* 23, 60–87. doi: 10.33423/jop.v23i3.6485

Mohammed, S., and Ringseis, E. (2001). Cognitive diversity and consensus in group decision making: the role of inputs, processes, and outcomes. *Organ. Behav. Hum. Decis. Process.* 85, 310–335. doi: 10.1006/obhd.2000.2943

Mohammed, S., and Schillinger, D. (2022). Translating time-based research into team interventions: An actionable, evidence-based approach. *J. Clin. Transl. Sci.* 6:e2. doi: 10.1017/cts.2021.874

Nam, C. S., Lyons, J. B., Hwang, H.-S., and Kim, S. (2009). The process of team communication in multi-cultural contexts: an empirical study using Bales' interaction process analysis (IPA). *Int. J. Ind. Ergon.* 39, 771–782. doi: 10.1016/j.ergon.2009.03.004

Narayan, S., Sidhu, J. S., and Volberda, H. W. (2021). From attention to action: the influence of cognitive and ideological diversity in top management teams on business model innovation. *J. Manage. Stud.* 58, 2082–2110. doi: 10.1111/joms.12668

Noy, S., and Zhang, W. (2023). Experimental evidence on the productivity effects of generative artificial intelligence. *Science* 381, 187–192. doi: 10.1126/science.adh2586

O'neill, T., McNeese, N., Barron, A., and Schelble, B. (2022). Human–autonomy teaming: a review and analysis of the empirical literature. *Hum. Factors* 64, 904–938. doi: 10.1177/0018720820960865

O'Neill, T. A., Flathmann, C., McNeese, N. J., and Salas, E. (2023). Human-autonomy teaming: need for a guiding team-based framework? *Comput. Hum. Behav.* 146:107762. doi: 10.1016/j.chb.2023.107762

Patrício, L., and Franco, M. (2022). A systematic literature review about team diversity and team performance: future lines of investigation. *Adm. Sci.* 12:31. doi: 10.3390/admsci12010031

Pelled, L. H., Eisenhardt, K. M., and Xin, K. R. (1999). Exploring the black box: an analysis of work group diversity, conflict and performance. *Admin. Sci. Q.* 44, 1–28. doi: 10.2307/2667029

Pieterse, A. N., Van Knippenberg, D., and van Ginkel, W. P. (2011). Diversity in goal orientation, team reflexivity, and team performance. *Organ. Behav. Hum. Decis. Process.* 114, 153–164. doi: 10.1016/j.obhdp.2010.11.003

Qi, M., Armstrong, S. J., Yang, Z., and Li, X. (2022). Cognitive diversity and team creativity: effects of demographic faultlines, subgroup imbalance and information elaboration. *J. Bus. Res.* 139, 819–830. doi: 10.1016/j.jbusres.2021.10.027

Raisch, S., and Krakowski, S. (2021). Artificial intelligence and management: the automation–augmentation paradox. *Acad. Manag. Rev.* 46, 192–210. doi: 10.5465/amr.2018.0072

Ridgeway, C. L. (2001). Gender, status, and leadership. *J. Soc. Issues* 57, 637–655. doi: 10.1111/0022-4537.00233

Rogers, Y., Lim, Y.-k., Hazlewood, W. R., and Marshall, P. (2009). Equal opportunities: do shareable interfaces promote more group participation than single user displays? *Hum. Comput. Interact.* 24, 79–116. doi: 10.1080/07370020902739379

Sauer, J., Felsing, T., Franke, H., and Rüttinger, B. (2006). Cognitive diversity and team performance in a complex multiple task environment. *Ergonomics* 49, 934–954. doi: 10.1080/00140130600577502

Schumpe, B. M., Bélanger, J. J., Nisa, C. F., and Pierro, A. (2023). When diversity leads to closed-mindedness: cognitive factors explain the effects of perceived diversity. *Curr. Psychol.* 42, 9324–9338. doi: 10.1007/s12144-021-02184-7

Sebo, S. S., Traeger, M., Jung, M., and Scassellati, B. (2018). The ripple effects of vulnerability: the effects of a robot's vulnerable behavior on trust in human-robot teams. 2018 13th ACM/IEEE International Conference on Human-Robot Interaction (HRI).

Shin, S. J., Kim, T. Y., Lee, J. Y., and Bian, L. (2012). Cognitive team diversity and individual team member creativity: a cross-level interaction. *Acad. Manag. J.* 55, 197–212. doi: 10.5465/amj.2010.0270

Simaremare, M. E., Pardede, C., Tampubolon, I. N., Simangunsong, D. A., and Manurung, P. E. (2024). The penetration of generative ai in higher education: a survey. 2024 IEEE Integrated STEM Education Conference (ISEC).

Simons, S. M., and Rowland, K. N. (2011). Diversity and its impact on organizational performance: the influence of diversity constructions on expectations and outcomes. *J. Technol. Manag. Innov.* 6, 171–183. doi: 10.4067/S0718-27242011000300013

Soukup, T., Lamb, B. W., Shah, N. J., Morbi, A., Bali, A., Asher, V., et al. (2020). Relationships between communication, time pressure, workload, task complexity, logistical issues and group composition in transdisciplinary teams: a prospective observational study across 822 cancer cases. *Front. Commun.* 5:583294. doi: 10.3389/fcomm.2020.583294

Srikanth, K., Harvey, S., and Peterson, R. (2016). A dynamic perspective on diverse teams: moving from the dual-process model to a dynamic coordination-based model of diverse team performance. *Acad. Manage. Ann.* 10, 453–493. doi: 10.5465/19416520.2016.1120973

Stasser, G., and Titus, W. (1985). Pooling of unshared information in group decision making: biased information sampling during discussion. *J. Pers. Soc. Psychol.* 48, 1467–1478. doi: 10.1037/0022-3514.48.6.1467

Stasser, G., and Titus, W. (2003). Hidden profiles: a brief history. *Psychol. Inq.* 14, 304–313. doi: 10.1207/S15327965PLI1403&4_21

Steiner, I. D. (1972). *Group process and productivity*. New York, NY: Academic press.

Stone, P., Kaminka, G., Kraus, S., and Rosenschein, J. (2010). *Ad hoc* autonomous agent teams: collaboration without pre-coordination. *Proc. AAAI Conf. Artif. Intell.* 24, 1504–1509. doi: 10.1609/aaai.v24i1.7529

Team, K., Du, A., Yin, B., Xing, B., Qu, B., Wang, B., et al. (2025). Kimi-VL Technical Report. *arXiv:2504.07491*. doi: 10.48550/arXiv.2504.07491

Van Knippenberg, D., and Schippers, M. C. (2007). Work group diversity. *Annu. Rev. Psychol.* 58, 515–541. doi: 10.1146/annurev.psych.58.110405.085546

Wang, X.-H. F., Kim, T.-Y., and Lee, D.-R. (2016). Cognitive diversity and team creativity: effects of team intrinsic motivation and transformational leadership. *J. Bus. Res.* 69, 3231–3239. doi: 10.1016/j.jbusres.2016.02.026

Ward, A. F. (2013). Supernormal: how the internet is changing our memories and our minds. *Psychol. Inq.* 24, 341–348. doi: 10.1080/1047840X.2013.850148

Zhang, R., Duan, W., Flathmann, C., McNeese, N., Freeman, G., and Williams, A. (2023). Investigating AI teammate communication strategies and their impact in human-AI teams for effective teamwork. *Proc. ACM Hum. Comput. Interact.* 7, 1–31. doi: 10.1145/3610072

Zhang, Z., Sun, B., and An, P. (2025). Breaking barriers or building dependency? Exploring Team-LLM collaboration in AI-infused classroom debate. *arXiv:2501.09165*. doi: 10.48550/arXiv.2501.09165

Zvelebilova, J., Savage, S., and Riedl, C. (2024). Collective attention in human-AI teams. *arXiv:2407.17489*. doi: 10.48550/arXiv.2407.17489

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 15 September 2025
DOI 10.3389/fpsyg.2025.1607274

Check for updates

OPEN ACCESS

EDITED BY
Ivan Wen,
University of Hawaii at Manoa, United States

REVIEWED BY
Wahyu Rahardjo,
Gunadarma University, Indonesia
Ibrahim Bozaci,
Kirikkale University, Türkiye

*CORRESPONDENCE
Wenjing Yu
✉ 280295388@qq.com

RECEIVED 07 April 2025
ACCEPTED 11 August 2025
PUBLISHED 15 September 2025

CITATION
Yu W and Chang X (2025) Exploration of factors of digital photo hoarding behavior among university students and the mediating role of emotional attachment and fear of missing out. *Front. Psychol.* 16:1607274.
doi: 10.3389/fpsyg.2025.1607274

Exploration of factors of digital photo hoarding behavior among university students and the mediating role of emotional attachment and fear of missing out

Wenjing Yu* and Xueqin Chang

Department of Electronic and Information Engineering, Bozhou University, Bozhou, Anhui, China

With the widespread use of digital technology and devices, college students are prone to hoarding digital photos. Based on the SOR model, this study conducted a survey of 294 college students and used partial least squares structural equation modeling (PLS-SEM) to study the factors of digital photo hoarding among college students, as well as the mediating effects of emotional attachment and fear of missing on the relationship between various factors and digital photo hoarding behavior. The results revealed that emotional attachment, fear of missing out, interpersonal influence, life demand, and technological progress are important influencing factors for college students' digital photo hoarding behavior. In addition, Emotional attachment mediates the relationship between emotional needs, interpersonal influence, and technological progress with digital photo hoarding behaviors. Fear of missing out mediates the association between emotional needs, interpersonal influence, and technological progress, and digital photo hoarding behavior. Finally, we discuss the implication, limitations, and directions for future research and conclusion of this work.

KEYWORDS

digital photo hoarding, fear of missing out, emotional attachment, SOR, PLS-SEM

Introduction

According to Photutorial statistics, by 2024, it is expected that 1.94 trillion photos will be taken globally, with 5.3 billion photos taken every day, or 61,400 photos per second. There are approximately 14.3 trillion existing photos, and photos taken by smartphones account for 94% of all photos. Google Image Search can search about 136 billion pictures, 14 billion pictures are shared every day on social media, and Americans take 20 pictures every day on average (Agarwal et al., 2024). With the reduction of digital storage costs and the continuous expansion of storage capacity, as well as the enhancement of digital shooting and editing tools, people are hoarding more and more photos on devices such as mobile phones, hard drives, and cloud drives, and are unwilling to organize or delete them. A study shows that a 47 year old man takes about 1,000 photos every day and saves them all. Although he never looks at or uses these photos, he believes they will be useful in the future. Organizing these photos left the man very frustrated and time-consuming, taking 3–5 h a day, seriously affecting his normal life (Bozaci and Gökdeniz, 2020). College students are active users of social media and an important group for hoarding digital photos. The study found that among 2,204 Chinese college students, 32.71% have hoarding behaviors (Zheng and Liu, 2020). The digital asset that college students hoard the most is photos, and the one they are least willing to delete is also photos. Photos are the main factor causing digital chaos (Broz, 2024). Hoarding digital photos not only causes digital chaos, but may also affect individual work efficiency, bring pressure and anxiety to hoarders, and even trigger cybersecurity issues (Chao and Li, 2023; Wu and Li, 2021). College students lack information literacy and organizational management skills. Studying the hoarding behavior of digital photos among college students can help them manage digital photos correctly, develop healthy digital habits, and avoid the negative effects of digital photo hoarding.

The current study

Van Bennekom first proposed the concept of digital hoarding, which he believed referred to the accumulation and chaos of digital files, as well as the difficulty of deleting them (Van Bennekom et al., 2015). Subsequently, many scholars have conducted research on digital hoarding. First, the negative impact of digital hoarding behavior. The behavior of digital hoarding will have a certain impact on computer science, psychology, and organizational science, causing problems in information security, information ethics, intellectual property, and so on (Guo et al., 2020; Zhao, 2020, 2025; Xu and Zhang, 2023). Digital hoarding is limited, and the more content is hoarded, the stronger the sense of loss caused by not hoarding content (Schüll, 2018). Digital hoarding behavior can affect individuals' work efficiency, increase psychological pressure and anxiety, and cause network problems (Sweeten et al., 2018; Zhao, 2022). And it will have a certain impact on an individual's cognition, emotions, and behavior. Second, development of a digital hoarding behavior scale. Neave et al. designed a new digital behaviors questionnaire (DBQ), including digital hoarding questionnaire (DHQ) and digital behaviors in the workplace questionnaire (DBWQ). The questionnaire mainly measures individuals' digital hoarding behavior during work (Jia et al., 2022). Based on the context of localization in China, some scholars developed a digital hoarding behavior scale that is tailored to individual characteristics in China (Kirk and Sellen, 2010; Guo et al., 2020; Wu et al., 2021a). Bozaci and Gökdeniz developed a digital photo hoarding behavior scale for individuals who hoard digital photos. Third, research on the influencing factors of digital hoarding behavior. Different scholars have studied the digital hoarding behavior of different individuals. The articles studied the influencing factors of college students' digital hoarding behavior (Wang et al., 2022; Jia et al., 2022; Zhang and He, 2023; Chao and Li, 2023; Guo, 2023; Oravec, 2018). The articles studied the influencing factors of digital hoarding behavior among social media users (Liu and Jia, 2023; Zhang and Liu, 2024a; Zhu and Jiang, 2024).

Digital photos are a type of digital content. Although there have been some studies on digital hoarding behavior at home and abroad, the granularity of research from the perspective of digital hoarding content is relatively coarse, and there is no distinction between digital hoarding content. Digital photos have the maximum share of hoarded digital content (Bozaci and Gökdeniz, 2020). With the increase in the capacity of data storage devices and the reduction in costs, as well as the upgrading of digital photo shooting tools, college students tend to use digital photos to record the little things in their lives. This article specifically studies the factors of digital photo hoarding among college students, as well as the mediating effects of emotional attachment and fear of missing on the relationship between various factors and digital photo hoarding behavior, which is of great significance for healthy digital content management of college students.

Research model and hypotheses

Research model

The SOR theory is the Stimulus-Organism-Response (S-O-R) model put forward by Mehrabian and Russell (1974). They believe that behavior is a response made by an individual's psychology influenced by external stimuli and then by the influence of psychology. In this model, the stimulus (S) represents the physical or non-physical stimuli that an individual receives, including those from the external environment, technological progress, and so on. The organism (O) represents the internal states such as cognition and emotion that an individual generates in response to the stimuli. The response (R) represents the behaviors that an individual exhibits after being stimulated. The behavior of digital photo hoarding is a reaction made by an individual due to factors such as the external environment, and it also involves changes in an individual's emotions and cognition. The SOR theory can well construct the relationships among external stimuli, an individual's internal states and behaviors, so it is applicable to the research on the behavior of digital photo hoarding.

The S-O-R model constitutes a causal chain of external stimuli, user cognition and behavior, and provides a detailed interpretation of the predictive effect of external stimuli on users' emotional responses and subsequent behaviors (Xu et al., 2025). Based on the SOR theory, a model of the influencing factors for college students' digital photo hoarding behavior is constructed, as shown in Figure 1. Among them, learning needs (LN), life demand (LD), emotional needs (EN), information overload (IO), interpersonal influence (II), and technological progress (TP) are regarded as the stimulus (S), emotional attachment (EA) and the fear of missing out (FoMO) are taken as the organism (O), and the digital photo hoarding behavior (DPHB) is considered as the response (R).This theoretical model mainly includes the following two causal relationships: (1) The internal or external stimuli (S) that college students receive directly affect their behavior of hoarding digital photos; (2) The stimuli (S) received by college students have an impact on their digital photo hoarding behavior through the mediating role of organic (O) emotional attachment and fear of missing out.

Hypotheses formation

Learning needs

The learning needs refer to the behavior of college students accumulating photos in order to increase their knowledge reserves or to cope with college assignments, exams, and other such activities. Digital picture hoarding enables people to save comprehensive and well-organized collections of images for various

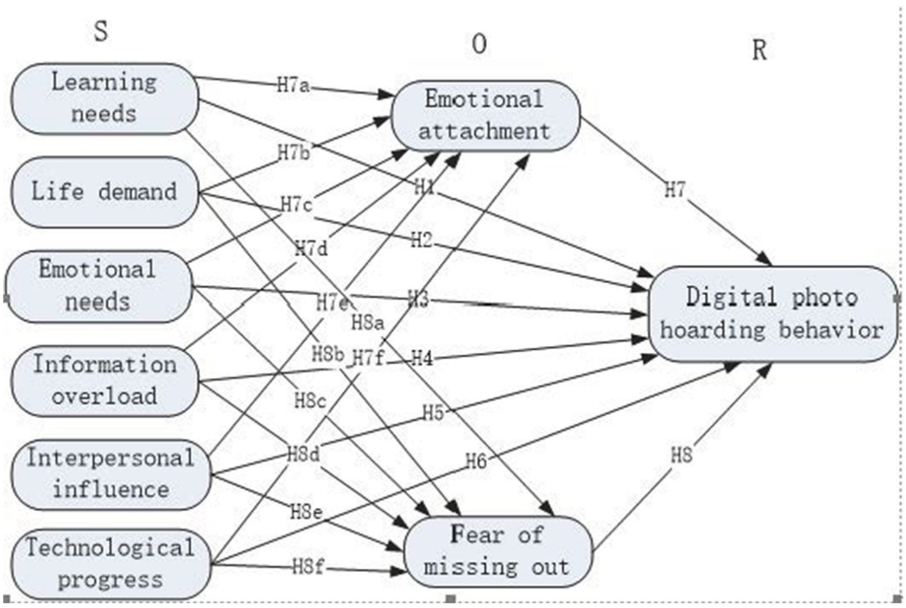


FIGURE 1
Theoretical model of the influencing factors for college students' digital photo hoarding behavior.

uses, including documentation, study, narrative, and private preservation (Liu et al., 2024). Users' personal needs can make them emotionally attached to data, which in turn affects their behavior. Hoarding data is mainly for academic research, seeking inspiration, and acquiring knowledge (Lu, 2023). Academic demands are an important factor for college students to develop digital hoarding behavior (Lu, 2024; Dai et al., 2024; Liu and Jia, 2023; He and Lin, 2025). Therefore, we proposed the following hypothesis:

H1: Learning needs has a significant positive effect on digital photo hoarding behavior.

Life demand

Besides hoarding digital photos due to learning needs, college students may also accumulate digital photos for security and livelihood guarantees. Sweeten et al. believes that future use as evidence is one of the motivations for digital hoarding behavior (Sweeten et al., 2018; Alquista and Baumeister, 2018). For example, they habitually back up data for fear of losing files, take screenshots of shopping and courier information and save them for easy checking at any time, and take pictures of personal identification documents and store them for reference when needed. Anaza and Nowlin believe that individuals tend to hoard important information in order to maintain their own advantages and enhance their competitiveness (Anaza and Nowlin, 2017). Therefore, we proposed the following hypothesis:

H2: Life demand has a significant positive effect on digital photo hoarding behavior.

Emotional needs

Individual hoarding behavior is associated with seven beliefs: remembering the past, defining the self, preventing forgetting, fulfilling responsibilities, building a family, maintaining connections with the past, and respecting those who care about us (Luxon et al., 2019). Grisham et al. believes that separation anxiety, uncertainty, interpersonal relationships, and perceived needs can affect data hoarding behavior (Grisham et al., 2019). Some college students hoard digital photos to satisfy their emotional needs. Viewing images of pleasant events, celebrations, and happy times can arouse positive feelings like joy, satisfaction, and appreciation, adding to a feeling of general wellbeing. Individuals hoard digital photos for nostalgia (Zheng and Liu, 2020; Feng, 2022; Fu et al., 2015). Digital photo hoarding can offer a therapeutic avenue for self-reflection, emotional expression, and mental health. Looking through one's digital photo collection can be a soothing and calming hobby. Butcher believes that individuals at work hoard data to gain a sense of security (Butcher, 1995). Therefore, we proposed the following hypothesis:

H3: Emotional needs have a significant positive effect on digital photo hoarding behavior.

Information overload

In the era of big data, an overwhelming amount of information is flooding in. Faced with the vast and diverse array of information, college students may accumulate a large number of digital photos due to their inability to organize and process the information they encounter, and they may also choose to store all information out of fear of missing out on important details. When there is too much information, it is difficult to judge the true value of the information. People often increase the frequency of using social media for fear of missing important information (Guo and Peng, 2025; Sun, 2023). There exists a positive association between information overload and the DHB exhibited by college students (Neave et al., 2019). Therefore, we proposed the following hypothesis:

H4: Information overload has a significant positive effect on digital photo hoarding behavior.

Interpersonal influence

Interpersonal influence refers to the impact that the environment and people around college students have on them. College students will share interesting images or videos they see on social media with classmates, relatives, and friends. College students enjoy the satisfaction that comes from social interaction when they share the digital photos they have stored. Digital photo hoarding is a helpful tool for improving relationships and is far more than just a habit of collecting photos. Users' social relationships and traditional cultural concepts, among others, can all have an impact on their emotions and behaviors (Lu, 2023). The act of gathering digital images can be a gratifying and relationship-enhancing activity (Agarwal et al., 2024). Additionally, research has found that upward social comparison has a positive impact on digital hoarding behavior (Wang et al., 2023; Liu and Jia, 2023). Therefore, we proposed the following hypothesis:

H5: Interpersonal influence has a significant positive effect on digital photo hoarding behavior.

Technological progress

Technological advancements have provided both software and hardware support for college students' digital photo hoarding. The upgrading of photography equipment and the continuous expansion of storage capacities for photos have led college students to not easily delete their favorite photos, and also encourage them to store a large number of digital photos to avoid missing important information (Agarwal et al., 2024). Technical support is one of the fundamental factors that enhance users' attachment to an App (Jin and Hou, 2022). External storage devices, application platforms, and network environments can all have an impact on users' emotions and behaviors (Lu, 2023). Therefore, we proposed the following hypothesis:

H6: Technological progress has a significant positive effect on digital photo hoarding behavior.

Mediating effect of emotional attachment

Emotional attachment has a significant impact on digital hoarding behavior (Luxon et al., 2019; Zhang and Liu, 2024b; Wu et al., 2021b). College students may develop emotional attachments to certain things due to their studies, life, and emotional experiences, and rely on technological support to engage in digital photo hoarding behavior. The emotional and personal meaning people attach to their digital photo collections is at the heart of the sentimental value of digital photo hoarding (Agarwal et al., 2024). Emotional attachment plays a mediating role in the impact of personal needs, personal habits, data characteristics, social influence, technical support, and data literacy on data hoarding behavior (Lu, 2023). Therefore, we proposed the following hypothesis:

H7: Emotional attachment mediates the association between (a) learning needs, (b) life demand, (c) emotional needs, (d) information overload, (e) interpersonal influence, (f) technological progress and digital photo hoarding behavior.

Mediating effect of fear of missing out

The fear of missing out (FoMO) is an important internal factor leading to digital photo hoarding among college students. College students may engage in digital photo hoarding behavior due to their academic, lifestyle, and emotional needs, while technological support, interpersonal influence, and information overload can exacerbate this behavior. The fear of missing out mediates the impact of upward social comparison on digital hoarding behavior (Wang et al., 2023; Liu and Jia, 2023). Therefore, we proposed the following hypothesis:

H8: Fear of missing out mediates the association between (a) learning needs, (b) life demand, (c) emotional needs, (d) information overload, (e) interpersonal influence, (f) technological progress and digital photo hoarding behavior.

Research method

Survey development and data collection

Referencing existing digital hoarding scales both domestically and internationally, and combining semi-structured interviews, the items for this questionnaire survey were ultimately determined after a preliminary research. The questionnaire uses a Likert five-point scale (ranging from 1 to 5, representing "strongly disagree," "disagree," "neutral," "agree," and "strongly agree") to measure the respondents' level of agreement with the items.

The questionnaire is divided into two parts. The first part mainly collects basic information about the respondents, including gender, grade, etc. The second part mainly investigates the factors influencing digital photo hoarding, including 9 observed variables: learning needs (LN), life demand (LD), emotional needs (EN), information overload (IO), interpersonal influence (II), technological progress (TP), emotional attachment (EA), fear of missing out (FoMO), and digital photo hoarding behavior (DPHB). The items for each observed variable and their reference sources are shown in Appendix 1.

Data collection was primarily conducted online, using QuestionStar to create the finalized questionnaire, which was then distributed to college students. After excluding invalid questionnaires, a total of 294 questionnaires were collected, with a response rate of 98%.

Results

Measurement model

The assessment of the measurement model encompassed an evaluation of its reliability, convergent validity, and discriminant validity. The reliability and validity of the measurement model were examined using the SmartPLS 4.0, as follows:

Reliability

Conducting reliability testing on the questionnaire can reveal the level of consistency. The measurement indicators include Cronbach's Alpha coefficient and Composite Reliability (CR). When Cronbach's Alpha is between 0.7 and 0.8, it

TABLE 1 The reliability of the measurement.

Constructs	Items	Factor loading	Cronbach's alpha	CR	AVE	VIF
DPHB	DPHB1	0.765	0.895	0.920	0.657	1.931
	DPHB2	0.789				2.026
	DPHB3	0.852				2.735
	DPHB4	0.852				3.158
	DPHB5	0.849				3.401
	DPHB6	0.749				1.905
EA	EA1	0.880	0.920	0.944	0.807	2.786
	EA2	0.919				3.791
	EA3	0.881				2.746
	EA4	0.912				3.454
EN	EN1	0.910	0.944	0.957	0.818	4.809
	EN2	0.898				4.255
	EN3	0.916				4.630
	EN4	0.894				3.752
	EN5	0.904				3.803
FOMO	FoMO1	0.895	0.883	0.919	0.740	2.940
	FoMO2	0.832				1.976
	FoMO3	0.855				2.212
	FoMO4	0.859				2.392
II	II1	0.749	0.753	0.859	0.671	1.344
	II2	0.868				1.903
	II3	0.836				1.677
IO	IO1	0.897	0.896	0.935	0.828	3.077
	IO2	0.890				2.328
	IO3	0.943				4.052
LN	LN1	0.858	0.725	0.878	0.783	1.478
	LN2	0.910				1.478
LD	LD1	0.932	0.839	0.925	0.861	2.091
	LD2	0.923				2.091
TP	TP1	0.827	0.899	0.926	0.714	2.187
	TP2	0.796				2.021
	TP3	0.859				2.617
	TP4	0.890				3.334
	TP5	0.850				2.649

indicates that the overall reliability of the questionnaire meets the requirements. When Cronbach's Alpha is >0.8, it indicates that the overall reliability of the questionnaire is good (Zaremohzzabieh et al., 2024). When Composite Reliability is >0.7, it indicates that the composite reliability of the questionnaire is good. Table 1 presents the Cronbach's Alpha, Composite Reliability, and Average Variance Extracted (AVE) for the questionnaire. From Table 1, it can be seen that the Cronbach's Alpha for all variables is >0.7, with 7 out of 8 variables having a Cronbach's Alpha coefficient above 0.8, and the Composite Reliability is >0.8 for all variables, indicating that the questionnaire has good reliability.

To assess multicollinearity, we also conducted a check on the Variance Inflation Factor (VIF). As shown in Table 1, all VIF values are below the recommended threshold of 5, confirming that there is no multicollinearity in the research model.

TABLE 2 Square root of construct's AVE and its correlation with any other construct.

Constructs	DPHB	EA	EN	FoMO	II	IO	LN	LD	TP
DPHB	**0.811**								
EA	0.749	**0.898**							
EN	0.517	0.740	**0.904**						
FoMO	0.782	0.832	0.664	**0.860**					
II	0.638	0.684	0.693	0.720	**0.819**				
IO	0.537	0.648	0.772	0.666	0.756	**0.910**			
LN	0.434	0.545	0.669	0.580	0.555	0.678	**0.885**		
LD	0.445	0.641	0.838	0.587	0.645	0.785	0.712	**0.928**	
TP	0.622	0.697	0.731	0.716	0.737	0.797	0.660	0.737	**0.845**

Bold values indicates the AVE >0.5.

Validity

Validity assessment can reveal the effectiveness of a questionnaire. The indicators of measurement validity include convergent validity and discriminant validity. Convergent validity is measured by the Average Variance Extracted (AVE). When the AVE is >0.5, it indicates good convergent validity (Zaremohzzabieh et al., 2024). Discriminant validity can be measured using the Fornell-Larcker Criterion and cross-loadings. As shown in Table 2, the values on the diagonal represent the square root of the AVE for each variable; the squared root of the AVE for each construct is greater than its correlation coefficients with other constructs. The item loadings of each construct are significantly higher than the cross-loadings of other constructs (Due to space constraints, the cross-loadings are not attached and can be requested from the authors upon request.). Therefore, the questionnaire has good discriminant validity.

Structural model

With the Bootstrapping in SmartPLS 4.0, the hypothesis testing results shown in Tables 3–5 were obtained.

Total effect

The total effect measures the entire influence of one variable on another, including the direct effect (represented by the path coefficient) and the indirect effect. It can comprehensively assess the importance of one variable to another, especially in complex models where there are mediator variables. It helps to understand how one variable influences another through multiple pathways. As shown in Table 3, emotional attachment, fear of missing out, interpersonal influence, life demand, and technological progress are important influencing factors for college students' digital photo hoarding behavior.

TABLE 3 Total effect.

Path	Coefficients	*T-values*	*P-values*	Result
EA -> DPHB	0.392***	4.567	0.000	Significant
EN -> DPHB	0.137	1.428	0.153	Insignificant
EN -> EA	0.484***	5.397	0.000	Significant
EN -> FoMO	0.243	2.757	0.006	Significant
FoMO-> DPHB	0.440***	5.146	0.000	Significant
II -> DPHB	0.430***	4.209	0.000	Significant
II -> EA	0.278**	3.402	0.001	Significant
II -> FoMO	0.381***	5.296	0.000	Significant
IO -> DPHB	−0.030	0.282	0.778	Insignificant
IO -> EA	−0.059	0.663	0.508	Insignificant
IO -> FoMO	0.036	0.388	0.698	Insignificant
LD -> DPHB	−0.182*	2.007	0.045	Significant
LD -> EA	−0.039	0.439	0.661	Insignificant
LD -> FoMO	−0.155	1.822	0.068	Insignificant
LN -> DPHB	0.063	1.033	0.302	Insignificant
LN -> EA	0.022	0.392	0.695	Insignificant
LN -> FoMO	0.152	1.851	0.064	Insignificant
TP -> DPHB	0.300***	3.622	0.000	Significant
TP -> EA	0.177*	2.116	0.034	Significant
TP -> FoMO	0.220**	2.744	0.006	Significant

*indicates $p < 0.05$; **indicates $p < 0.01$; ***indicates $p < 0.001$.

Path coefficients and specific indirect effects

The path coefficient represents the strength of the direct causal relationship between variables. In a path model, the arrow from an independent variable (predictor variable) to a dependent variable (predicted variable) represents a hypothesized causal connection, and the path coefficient is the quantification of the strength of this connection. As shown in Table 4, the results show that emotional attachment has a significant positive effect on digital photo hoarding. Emotional need, interpersonal influence, and technological progress have significant positive effects on emotional attachment. Hence, H7, H7c, H7e, H7f are supported. Fear of missing out has a significant positive effect on digital photo hoarding. Emotional need, interpersonal influence, and

TABLE 4 Parameter estimates for the path model predicting digital photo hoarding behavior.

Hypothesis	Path	Coefficients	*T*-values	*P*-values	Supported
H1	LN -> DPHB	−0.013	0.198	0.843	No
H2	LD -> DPHB	−0.098	1.281	0.200	No
H3	EN -> DPHB	−0.159	2.343	0.019	No
H4	IO -> DPHB	−0.023	0.256	0.798	No
H5	II -> DPHB	0.154	1.760	0.078	No
H6	TP -> DPHB	0.133	1.721	0.085	No
H7	EA -> DPHB	0.392***	4.567	0.000	Yes
H7a	LN -> EA	0.022	0.392	0.695	No
H7b	LD -> EA	−0.039	0.439	0.661	No
H7c	EN -> EA	0.484***	5.397	0.000	Yes
H7d	IO -> EA	−0.059	0.663	0.508	No
H7e	II -> EA	0.278**	3.402	0.001	Yes
H7f	TP -> EA	0.177*	2.116	0.034	Yes
H8	FoMO-> DPHB	0.440***	5.146	0.000	Yes
H8a	LN -> FoMO	0.152	1.851	0.064	No
H8b	LD -> FoMO	−0.155	1.822	0.068	No
H8c	EN -> FoMO	0.243**	2.757	0.006	Yes
H8d	IO -> FoMO	0.036	0.388	0.698	No
H8e	II -> FoMO	0.381***	5.296	0.000	Yes
H8f	TP -> FoMO	0.220**	2.744	0.006	Yes

*indicates $p < 0.05$; **indicates $p < 0.01$; ***indicates $p < 0.001$.

technological progress have significant positive effect on fear of missing out. Hence, H8, H8c, H8e, H8f are supported.

As shown in Table 5 and Figure 2, both emotional attachment and fear of missing out mediate the relationship between emotional need, interpersonal influence, and technological progress with digital photo hoarding behavior.

Figure 2 is a theoretical model with path coefficients based on the structural equation model.

Discussion

The article examines the influencing factors affecting digital photo hoarding among college students and the mediating roles of emotional attachment and fear of missing out in it. Through a questionnaire survey and using structural equation modeling, it is found that interpersonal influence, life demand, technological progress, emotional attachment, and fear of missing out have a significant impact on digital photo hoarding behavior. Learning needs, emotional needs, information overload have no significant impact on the hoarding behavior of digital photos. Moreover, both emotional attachment and fear of missing out mediate the effects of emotional need, interpersonal influence, and technological progress on digital photo hoarding behavior.

College students are surrounded by classmates, relatives, and friends who influence their digital photo hoarding behavior. College students choose to hoard digital photos because of peer comparisons or in order to keep good memories with family and friends, which is consistent with previous research that upward comparisons have a significant effect on digital hoarding behavior (Liu and Jia, 2023) and that people choose not to delete photos in order to keep the good moments (Agarwal et al., 2024). Life demand also lead to digital photo hoarding, and college students will choose not to delete digital photos for a long time in order to keep evidence, credentials, etc. (Liu and Jia, 2023). Technological progress provides convenient conditions for digital photo hoarding, and the expansion of storage space, shrinking costs, and the convenience of cross-platform storage all provide conditions for college students to hoard digital photos, which is consistent with the results that perceived low price and perceived convenience have a positive effect on digital hoarding behavior (Vinoi et al., 2024). Learning needs, emotional needs and information overload have no direct impact on college students' behavior of hoarding digital photos. It is indicated that the hoarding behavior of digital photos among college students will not be affected by learning needs and information overload. However, emotional needs influence college students' behavior of hoarding digital photos by affecting emotional attachment and fear of missing out. The emotional and personal meaning people attach to their digital photo collections is at the heart of the sentimental value of digital photo hoarding. Gratification derived from digital photo attachment can strongly induce people to gather and store vast collections of digital photos (Agarwal et al., 2024).

TABLE 5 The mediation effects on digital photo hoarding behavior.

Path	Coefficients	*T-values*	*P-values*	Supported
EN -> EA -> DPHB	0.190**	3.260	0.001	Yes
II -> FoMO -> DPHB	0.167**	3.439	0.001	Yes
IO -> FoMO -> DPHB	0.016	0.370	0.712	No
II -> EA -> DPHB	0.109	2.789	0.005	Yes
LD -> FoMO -> DPHB	−0.068	1.779	0.075	No
IO -> EA -> DPHB	−0.023	0.647	0.518	No
LN -> FoMO -> DPHB	0.067	1.847	0.065	No
LD -> EA -> DPHB	−0.015	0.431	0.667	No
TP -> FoMO -> DPHB	0.097*	2.382	0.017	Yes
LN -> EA -> DPHB	0.009	0.394	0.694	No
TP -> EA -> DPHB	0.069*	1.893	0.048	Yes
EN -> FoMO -> DPHB	0.107*	2.539	0.011	Yes

*indicates $p < 0.05$; **indicates $p < 0.01$; ***indicates $p < 0.001$.

It was also found that emotional attachment and fear of missing out mediate the effects of emotional need, interpersonal influence, and technological progress on digital photo hoarding behaviors. Many scholars in the past have used emotional attachment and fear of missing out as mediators in the study of digital hoarding behaviors. For example, emotional attachment mediates the effects of personal habit, personal need, social influence, and technological support on digital hoarding behaviors (Lu, 2023). Fear of missing out mediates the effect of upward social comparison on digital hoarding behavior (Liu and Jia, 2023; Wang et al., 2023).

Implication

Theoretical implication

Although scholars at home and abroad have conducted research on topics related to digital hoarding, most of them have studied the influencing factors of digital hoarding behavior and the moderating and mediating effects of different factors on digital hoarding behavior. Fewer studies have been conducted on specific digital hoarding content. In the new media era, with the upgrading of storage devices, storage space, and photographic technology, digital photographs are the type of data with more digital content storage. Studying the factors influencing college students' digital photo hoarding behaviors and the mediating roles of emotional attachment and misplaced fear in them has several theoretical implications:

First, it provides a new research perspective for digital hoarding research, different groups hoard different types of digital content, and this study provides reference and reflection for refining digital hoarding content research.

Second, based on the SOR model, structural equation modeling is used to study the influencing factors of digital photo hoarding behavior and the mediating role of emotional attachment and fear of missing out, which provides theoretical and methodological reference for digital hoarding behavior research.

Third, not only study the influencing factors of digital hoarding behavior, but also analyze the mediating role played by multimediating variables in the influence of emotional need, interpersonal influence and technological progress on digital photo hoarding behavior.

Practical implications

This study has not only some theoretical significance, but also some practical significance. College students will engage in digital photo hoarding because of comparisons with classmates around them, as well as because of the demands of life, and will also be too lazy to delete photos simply because of the convenience and low cost of storage. The massive hoarding of digital photos may cause the leakage of information, and may also bring some anxiety and pressure to college students. Therefore, it is necessary to organize digital photos at regular intervals, and college students should store digital photos reasonably, make rational use of digital resources, and develop good digital habits.

Limitations and directions for future research

Although this study has certain contributions, there are some shortcomings. First of all, the study's research subjects are college students, which is not generalizable. Future studies should include research subjects from different backgrounds and cultures, either individuals or organizations. For example, the hoarding behavior of digital photos by postgraduate students, research institutions, data resource management departments, etc. Study the differences in digital hoarding behaviors among college students in different countries. Secondly, the study only analyzed the behavior of mediating variables on digital photo hoarding behavior, and did not involve the study of moderating variables, which can be studied in the future to investigate the influence

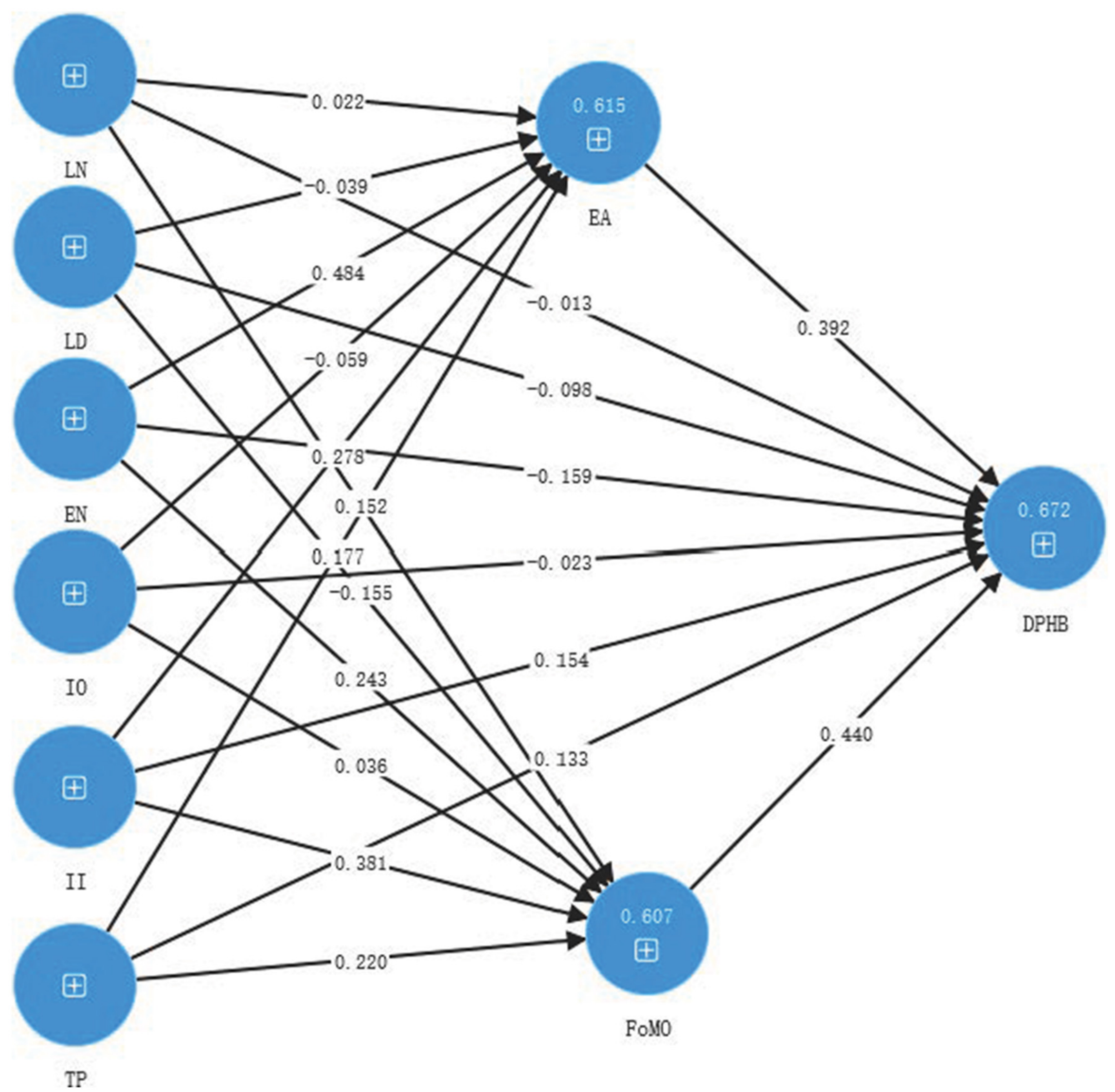


FIGURE 2
Results from the PLS model.

of moderating factors on digital photo hoarding behavior. For example, the moderating effect of conservative on the association between emotional attachment and digital hoarding behavior. Thirdly, there are certain errors in the data collected through questionnaires. For instance, the respondents deliberately choose the "socially expected answers" to conform to social norms, gain recognition from others, or avoid negative evaluations, rather than their true thoughts or behavioral tendencies. Or, due to factors such as self-awareness, motivation, or context, there may be deviations where the reported content does not match the actual behavior. Future research can combine objective data verification to improve the accuracy of source data. Fourthly, with the change of time and technology, the factors affecting the digital photo hoarding behavior of college students may change, and future research can explore other factors affecting the digital photo hoarding behavior on the basis of this study, and also study the relationship between digital photo hoarding behavior and physical hoarding behavior.

Conclusion

This article takes college students as the research object and studies their hoarding behavior of digital photos based on the SOR model. By using the structural equation model and conducting a survey among 294 college students, it was found that interpersonal influence, life demand, and technological progress have an important impact on college students' digital photo hoarding. Emotional attachment mediates the relationship between emotional needs, interpersonal influence, and technological progress with digital photo hoarding behaviors. Fear of missing out mediates the association between emotional needs, interpersonal influence, and technological progress, and digital photo hoarding behavior. Hence, H7, H7c, H7e, H7f, H8, H8c, H8e, H8f are supported. Although the research has certain limitations, it also makes certain theoretical and practical contributions, which not only broaden the research direction for the study of digital hoarding, but also help to guide college students to use digital content correctly, improve digital literacy, and develop good digital governance habits.

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.

Author contributions

WY: Writing – original draft, Writing – review & editing. XC: Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This work was supported by the Research Status, Implementation Challenges, and Optimization Strategies of Student Value added Evaluation [2024XJXM072]; Research on the Reform and Practice of Smart Curriculum Empowered by AI + Knowledge Graph: Taking E-commerce Organization and Operation as an Example [2024jyxm0499]; E-commerce New Professional Quality Improvement Project [2023xjzlts083].

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.

Generative AI statement

The author(s) declare that no Gen AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2025.1607274/full#supplementary-material

References

Agarwal, R., Mehrotra, A., Pant, M. K., Alzeiby, E. A., and Vishnoi, S. K. (2024). Digital photo hoarding in the online retail context: an in-depth qualitative investigation of retail consumers. *Int. J. Retail Distrib. Manag.* 78, 1–18. doi: 10.1016/j.jretconser.2024.103729

Alquista, J. L., and Baumeister, R. F. (2018). Foraging extends beyond food: hoarding of mental energy and information seeking in response to uncertainty. *Behav. Brain Sci.* 42, 17–18. doi: 10.1017/S0140525X18001838

Anaza, N. A., and Nowlin, E. L. (2017). What's mine is mine: a study of salesperson knowledge withholding and hoarding behavior. *Ind. Mark. Manag.* 64, 14–24. doi: 10.1016/j.indmarman.2017.03.007

Bozaci, I., and Gökdeniz, I. (2020). Development of a digital photo hoarding scale: a research with undergraduate students. *Manag. Sci. Lett.* 10, 2193–2200. doi: 10.5267/j.msl.2020.3.017

Broz, M. (2024). Photo statistics: how many photos are taken every day? *Photutorial.* Available online at: https://photutorial.com/photos-statistics/ (Accessed September 10, 2024).

Butcher, H. (1995). "Information overload in management and business," in *IEE Colloquium on Information Overload* (London: IET), 1/1–1/2.

Chao, Y., and Li, F. (2023). Formation mechanism of digital hoarding behavior among college students. *Encyclopedic Knowl.* 2023, 26–28.

Dai, B., Zhang, P. J., and Wang, Y. Q. (2024). Influencing factors and consequences of postgraduates' academic digital hoarding behavior: an exploratory study based on grounded theory. *Mod. Inf.* 44, 53–65. doi: 10.3969/j.issn.1008-0821.2024.04.006

Feng, Q. S. (2022). *The influence of loneliness on college students' hoarding behavior: The serial mediating role of anthropomorphism and object attachment* [Doctoral dissertation]. Yangtze University, Jingzhou, China.

Fu, Z. F., Xu, W., and Wang, J. P. (2015). The relationship between attachment anxiety, hoarding, and negative emotions: the mediating role of hoarding beliefs. *Chin. J. Clin. Psychol.* 23, 660–664. doi: 10.16128/j.cnki.1005-3611.2015.04.021

Grisham, J. R., Frost, R. O., Steketee, G., Kim, H. J., and Hood, S. D. (2019). The nature and treatment of compulsive hoarding: a review of the science and practice. *J. Obsess. Compuls. Relat. Disord.* 22:100255. doi: 10.1016/j.jocrd.2019.100255

Guo, H., Wei, X., Zhao, S., and Jin, T (2020). Revision of the Chinese version of digital hoarding behaviors questionnaire. *Chin. J. Clin. Psychol.* 28, 492–495. doi: 10.16128/j.cnki.1005-3611.2020.04.025

Guo, F. Y., and Peng, X. (2025). Excessive storage and difficulty in deletion: mechanism model and element relationship of digital hoarding behavior among youth groups——A qualitative study based on grounded theory. *Libr. Theory Pract.* 114–126.

Guo, Y. (2023). *The impact of perfectionist personality on digital hoarding* [Master's thesis]. Southwest University, Chongqing, China.

He, Y., and Lin, C. (2025). Research on influencing factors of postgraduates' academic digital hoarding behavior on social media platforms: an empirical analysis based on SEM and fsQCA. *Libr. Work Res.* 16–30+63.

Jia, M. X., Xu, Y. Q., and Zhao, Y. X. (2022). An exploratory study on digital hoarding behavior of college students: from the perspective of personal information management. *Libr. Inf. Serv.* 66, 74–88. doi: 10.13266/j.issn.0252-3116.2022.12.012

Jin, Y. C., and Hou, Q. Q. (2022). Analysis on aging-friendly transformation of audio apps from the perspective of attachment theory. *New Media Res.* 8, 28–32. doi: 10.3969/j.issn.2096-0360.2022.12.014

Kirk, D. S., and Sellen, A. (2010). On human remains: values and practice in the home archiving of cherished objects. *ACM Trans. Comput. Hum. Interaction* 17:10. doi: 10.1145/1806923.1806924

Liu, T. Y., and Jia, Y. (2023). The formation mechanism and outcome effects of digital hoarding behavior among youth groups: an analysis from the perspective of internal motivations. *Chinese Youth Stud.* 34, 93–100. doi: 10.19633/j.cnki.11-2579/d.2023.0025

Liu, Y., Chi, X., and Xin, X. (2024). Storing, not reading: Investigating the link between upward social comparison via social media and digital hoarding behavior in Chinese youth. *Psychol. Res. Behav. Manag.* 17, 5209–5224. doi: 10.2147/PRBM.S441859

Lu, J. L. (2024). *Factors influencing academic data hoarding behavior among university students* [Master's thesis]. Hebei University, Baoding, China.

Lu, J. Q. (2023). *A study on the influencing factors of graduate students' data hoarding behavior* [Doctoral dissertation]. Central China Normal University, Wuhan.

Luxon, A. M., Hamilton, C. E., Bates, S., and et al. (2019). Pinning our possessions: associations between digital hoarding and symptoms of hoarding disorder. *J. Obsess. Compuls. Relat. Disord.* 21, 60–68. doi: 10.1016/j.jocrd.2018.12.007

Mehrabian, A., and Russell, J. A. (1974). *An Approach to Environmental Psychology*. Cambridge, MA: MIT Press.

Neave, N., Briggs, P., McKellar, K., and Sillence, E. (2019). Digital hoarding behaviours: measurement and evaluation. *Comput. Hum. Behav.* 96, 72–77. doi: 10.1016/j.chb.2019.01.037

Oravec, A. J. (2018). Digital (or Virtual) hoarding: emerging implications of digital hoarding for computing, psychology, and organization science. *Int. J. Comput. Clin. Pract.* 3, 27–39. doi: 10.4018/IJCCP.2018010103

Schüll, D. N. (2018). Digital containment and its discontents. *Hist. Anthropol.* 29, 42–48. doi: 10.1080/02757206.2017.1397654

Sun, C. L. (2023). *A study on the influencing factors of digital hoarding behavior among social media users* [Doctoral dissertation]. Tianjin Normal University, Xiqing, China.

Sweeten, G., Sillence, E., and Neave, N. (2018). Digital hoarding behaviours: underlying motivations and potential negative consequences. *Comput. Hum. Behav.* 85, 54–60. doi: 10.1016/j.chb.2018.03.031

Van Bennekom, M. J., Jansen, A. M., and Olde Hartman, M. A. (2015). A case of digital hoarding. *BMJ Case Rep.* 2015:bcr2015210814. doi: 10.1136/bcr-2015-210814

Vinoi, N., Shankar, A., Khalil, A., Mehrotra, A., and Kumar, J. (2024). Holding on to your memories: factors influencing social media hoarding behaviour. *J. Retail. Consum. Serv.* 76, 1–12. doi: 10.1016/j.jretconser.2023.103617

Wang, H., Miao, P., Jia, H., and Lai, K. (2023). The dark side of upward social comparison for social media users: an investigation of fear of missing out and digital hoarding behavior. *Soc. Media Soc.* 9, 1–13. doi: 10.1177/20563051221150420

Wang, L., Du, T. Y., and Zhu, H. J. (2022). Formation mechanism of college students' data hoarding behavior in social media environment. *Inf. Stud.* 45, 22–29. doi: 10.16353/j.cnki.1000-7490.2022.01.004

Wu, X. Y., Huang, X., and Li, J. (2021a). Development and reliability and validity testing of the digital hoarding behavior scale. *Psychol. Technol. Appl.* 9, 116–125. doi: 10.16842/j.cnki.issn2095-5588.2021.02.007

Wu, X. Y., Huang, X., and Li, J. (2021b). The relationship between attachment anxiety and digital hoarding behavior: the mediating role of intolerance of uncertainty and difficulties in emotion regulation. *Chin. J. Clin. Psychol.* 29, 996–999. doi: 10.16128/j.cnki.1005-3611.2021.05.020

Wu, X. Y., and Li, J. (2021). "Digital possession" in the information age: digital hoarding and its related research. *Psychol. Sci.* 44, 800–806. doi: 10.16719/j.cnki.1671-6981.20210405

Xu, X. F., and Zhang, Y. (2023). Hobby of hoarding in the digital age: a review of research on digital hoarding behavior. *Southeast Commun.* 20, 100–103. doi: 10.13556/j.cnki.dncb.cn35-1274/j.2023.09.016

Xu, X. K., Yan, Y. J., and Gao, W. (2025). Based on the S-O-R model: research on the identification of digital hoarding behavior among graduate students. *Inf. Sci.* 43, 161–168. doi: 10.13833/j.issn.1007-7634.2025.01.017

Zaremohzzabieh, Z., Abdullah, H., Ahrari, S., Abdullah, R., and Md Nor, S. M. (2024). Exploration of vulnerability factors of digital hoarding behavior among university students and the moderating role of maladaptive perfectionism. *Digit. Health* 10, 1–16. doi: 10.1177/20552076241226962

Zhang, Y., and Liu, M. (2024b). Research on the mechanisms of digital hoarding behavior of mobile social media users from the perspective of phenomenological diagnostics. *Mod. Intell.* 44, 3–12+73. doi: 10.12154/j.qbzlgz.2024.02.006

Zhang, Y. F., and Liu, M. (2024a). Generative mechanism and guidance strategy of digital hoarding behavior of mobile social media users. *Inf. Doc. Serv.* 45, 48–56.

Zhang, Z., and He, W. (2023). Influencing factors and configurational paths of digital hoarding behavior among college students. *Inf. Stud.* 46, 108–114. doi: 10.16353/j.cnki.1000-7490.2023.01.013

Zhao, C. (2025). Representation, psychological motivation and mechanism analysis of digital hoarding behavior among youth groups in the digital age. *New Media Res.* 11, 75–78+83. doi: 10.16604/j.cnki.issn2096-0360.2025.09.005

Zhao, D. X. (2022). A review and prospect of research on individual digital hoarding behavior. *J. Intell.* 41:8. doi: 10.3969/j.issn.1002-1965.2022.08.026

Zhao, D. X. (2020). *A study on individual digital hoarding behavior in the big data environment* [Doctoral dissertation]. Wuhan University, Wuhan, China.

Zheng, Y. N., and Liu, D. X. (2020). Investigation on the phenomenon of hoarding and influencing factors among Chinese college students. *Chin. J. Public Health* 36:4. doi: 10.11847/zgggws1124563

Zhu, L. L., and Jiang, H. L. (2024). The practice of negative performativism: a study on digital hoarding behavior of youth groups on social platforms. *J. Int. Commun.* 46, 154–176. doi: 10.13495/j.cnki.cjjc.2024.06.004

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 17 September 2025
DOI 10.3389/fpsyg.2025.1669247

Check for updates

OPEN ACCESS

EDITED BY
Wenjing Chen,
Beijing University of Posts and Telecommunications (BUPT), China

REVIEWED BY
Jing He,
Kunming University, China
Liu Xuan,
Nanjing Tech University, China
Abdullah Aloqab,
Hunan University, China

*CORRESPONDENCE
Ling Huang
✉ huang@ynu.edu.cn
Yuping Zhao
✉ zhaoyupingbupt@163.com

[†]These authors have contributed equally to this work

RECEIVED 19 July 2025
ACCEPTED 20 August 2025
PUBLISHED 17 September 2025

CITATION
Huang L and Zhao Y (2025) The impact of AI literacy on work–life balance and job satisfaction among university faculty: a self-determination theory perspective. *Front. Psychol.* 16:1669247.
doi: 10.3389/fpsyg.2025.1669247

The impact of AI literacy on work–life balance and job satisfaction among university faculty: a self-determination theory perspective

Ling Huang[1*†] and Yuping Zhao[2*†]

[1]School of Business and Tourism Management, Postdoctoral Research Station in Business Administration, Yunnan University, Kunming, China, [2]General Education Center, Communication University of China, Beijing, China

Introduction: The emergence of artificial intelligence (AI) is transforming the nature of academic work, yet the role of AI literacy in supporting faculty well-being remains underexplored. This study investigates how AI literacy influences university faculty's work-life balance and job satisfaction through the satisfaction of three basic psychological needs.
Methods: Survey data were collected from 511 faculty members. Measures included AI literacy, perceived autonomy, perceived competence, perceived relatedness, work-life balance, job satisfaction, and technology acceptance. Statistical analyses examined the direct and indirect effects of AI literacy on faculty well-being.
Results: The findings indicate that AI literacy significantly enhances the satisfaction of autonomy, competence, and relatedness. These, in turn, promote greater work-life balance. Further analysis shows that only perceived autonomy directly predicts job satisfaction, while competence and relatedness influence job satisfaction indirectly through work-life balance. Technology acceptance was found to moderate the relationship between AI literacy and psychological need fulfillment.
Discussion: This study illuminates the psychological pathways through which AI literacy contributes to faculty well-being. It extends the application of Self-Determination Theory to technology-intensive academic settings and offers practical implications for designing AI literacy initiatives and faculty support strategies in higher education.

KEYWORDS
AI literacy, university faculty, self-determination theory, work–life balance, job satisfaction, technology acceptance

1 Introduction

The swift rise of artificial intelligence (AI) has reshaped how work is structured and performed in educational contexts (Zhang, 2023). In higher education, AI is increasingly embedded in curriculum design, instructional analytics, assessment feedback, and other core academic activities (Hwang et al., 2020). While these technologies enhance instructional efficiency, they also present new challenges, including evolving professional roles, continuous demands for upskilling, and increased risks of psychological strain and job burnout (Yu, 2024; Zhou J. S. et al., 2024). Faculty members must not only adapt to rapidly changing technologies but also manage elevated workloads and mounting psychological pressures. In this context, developing AI literacy—a composite of knowledge, attitudes, and competencies necessary to understand and apply AI tools—has become an essential skill set for academic professionals (Laupichler et al., 2022; Ng et al., 2023; Sperling et al., 2024).

Although existing research has begun to examine the relationship between AI literacy and educators' professional experiences, key psychological dimensions remain insufficiently explored. For instance, Hashem et al. (2024) found that higher levels of AI literacy improve teaching effectiveness and reduce burnout, though their emphasis was primarily on technological performance. Similarly, Bhojak et al. (2025) reported a positive correlation between AI proficiency and job satisfaction, but did not investigate the underlying psychological mechanisms. Zheng and Zhang (2025) noted that the increased use of educational technology may blur work–family boundaries; however, their analysis lacked AI-specific focus and failed to address psychological needs explicitly.

Despite these preliminary insights, there remains a paucity of systematic empirical research on how AI literacy shapes university faculty's psychological functioning, subjective well-being, and work–life balance (Ding et al., 2024). Both work–life balance and job satisfaction are crucial indicators of faculty well-being and are closely tied to mental and physical health, professional engagement, and long-term career sustainability (Landolfi et al., 2021).

Self-Determination Theory (SDT) provides a well-established framework for understanding these psychological processes. According to SDT, the satisfaction of three basic psychological needs—perceived autonomy, perceived competence, and perceived relatedness—is fundamental to intrinsic motivation, psychological well-being, and job satisfaction (Ryan and Deci, 2020). Prior studies suggest that satisfying these needs significantly influences individuals' experiences with emerging technologies (Shen and Cui, 2024). Moreover, faculty members' level of technology acceptance may moderate the extent to which AI literacy supports psychological need fulfillment (Pan, 2020).

Nonetheless, notable gaps persist. First, much of the literature focuses narrowly on performance-related outcomes, overlooking AI literacy's potential role in meeting faculty members' psychological needs (Bhojak et al., 2025; Xiao et al., 2025). Second, few studies have empirically validated the psychological mechanisms underpinning this relationship through the lens of SDT (Zhou J. S. et al., 2024; Zhou T. et al., 2024). Third, existing models often neglect the moderating influence of technology acceptance, limiting their explanatory power (Şimşek, 2025). Crucially, an integrative framework that brings together AI literacy, SDT, and technology acceptance is still lacking.

To address these gaps, the present study focuses on university faculty and investigates the following research questions: (1) Does AI literacy influence perceived autonomy, perceived competence, and perceived relatedness? (2) Do these psychological needs contribute to enhanced work–life balance and job satisfaction? (3) Does technology acceptance moderate these psychological pathways?

By integrating AI literacy with foundational psychological constructs, the study aims to identify key psychological constructs involved in faculty adaptation to the evolving demands of AI-mediated academic work. The findings offer theoretical insights and practical implications for promoting the sustainable and psychologically supportive adoption of AI technologies in higher education. This paper is organized into the following sections: theoretical background and literature review, research model and hypotheses, methodology, data analysis, and discussion.

2 Theoretical framework and literature review

2.1 AI literacy

To conceptualize artificial intelligence (AI) literacy in the context of higher education, a literature search and thematic analysis were conducted using the core search terms "AI literacy" and "teacher" in

TABLE 1 Summary of AI literacy conceptualizations and related research contributions.

Reference	Conceptual definition	Research method	Research contribution
Kelley and Wenzel (2025)	AI literacy refers to the capacity to engage with AI tools effectively and ethically, critically evaluate AI outputs, and flexibly adapt across diverse environments.	Qualitative Study	Conducted a systematic review of the multi-stage integration of generative AI in teacher education, distilling key practical experiences.
Ayanwale et al. (2024)	AI literacy is defined as a comprehensive ability to critically understand, appropriately apply, and actively engage with AI technologies across various contexts.	Quantitative Study	Examined the current status of AI literacy among pre-service teachers in Nigeria and proposed a localized teacher training framework, providing empirical evidence on technical and ethical dimensions.
Ning et al. (2025)	AI literacy encompasses the knowledge and application skills required for teachers to adapt to intelligent teaching environments, including perception, knowledge, skills, practical application, and ethics.	Quantitative Study	Developed and validated a measurement scale for teachers' AI literacy, enriching the structural dimensions and assessment tools for AI literacy.
Sperling et al. (2024)	AI literacy refers to teachers' integrated ability to combine AI knowledge, teaching skills, and ethical judgment in educational practice.	Qualitative Study	Conducted a systematic review of AI literacy in teacher education, identified research gaps, and proposed recommendations for in-situ AI literacy and ethical training.
Ozudogru and Durak (2025)	AI literacy is the ability to use AI products and tools and to assess their potential social and environmental impacts.	Quantitative Study	Developed and validated a structural path model linking pre-service teachers' AI readiness to innovation, perceived threats, and AI literacy.
Al-Abdullatif (2024)	AI literacy integrates AI-related knowledge, skills, teaching, and assessment competencies.	Quantitative Study	Empirically investigated the key drivers of teachers' adoption of generative AI, focusing on the relationships among Technology Acceptance, technological pedagogical content knowledge (TPACK) for intelligent technologies, AI literacy, and perceived trust.

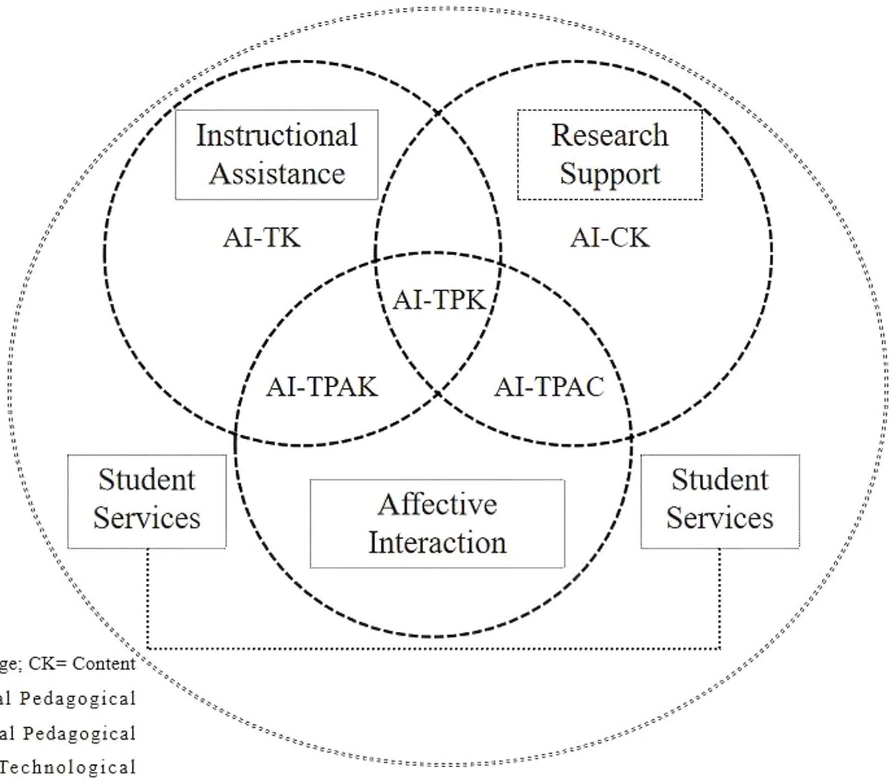
Source(s): Created by author.

the Web of Science database. The results, summarized in Table 1, reveal that definitions of AI literacy are diverse and continuously evolving. Some scholars define AI literacy as the capacity to engage with AI tools effectively and ethically, critically assess their outputs, and adapt to rapidly changing technological environments (Kelley and Wenzel, 2025; Ozudogru and Durak, 2025). Others emphasize its interactive and collaborative dimensions, suggesting that AI literacy involves active engagement with AI systems that goes beyond mere technical proficiency or tool operation (Ayanwale et al., 2024).

In higher education, faculty members serve as primary users of AI technologies in both teaching and research. Their AI literacy often manifests in complex, multidimensional ways (Laupichler et al., 2022). Drawing from prior literature and observed academic practices, faculty typically apply AI tools across four core domains: instructional assistance, research support, student services, and affective interaction (Yu, 2024; Zawacki-Richter et al., 2019). Their interactions with AI systems can be characterized through several common mechanisms (Şimşek, 2025; Kim, 2024), including: Information input–feedback, such as generating lesson content; Conflict resolution–trust building, for reconciling discrepancies between human and AI suggestions; Task delegation–cognitive offloading, involving the automation of repetitive or routine tasks; Decision support–human–machine collaboration, where AI contributes to complex academic analyses. These interaction patterns help operationalize how AI literacy manifests in faculty members' daily practices. As illustrated in Figure 1, AI literacy is distributed across intersecting domains of technological (TK), content (CK), and pedagogical (PK) knowledge, giving rise to integrated competencies such as AI-TPK, AI-TPAC, and AI-TPAK. This framework highlights the dynamic and interdisciplinary nature of AI literacy in higher education (see Figure 1).

An expanding body of research identifies AI literacy as a critical competency for instructional design, pedagogical decision-making, and student-centered learning. It plays an essential role in curriculum development, formative assessment, and personalized instruction (Al-Abdullatif, 2024; Ning et al., 2025). Several scholars further advocate for a broadened conceptualization of teacher AI literacy that integrates technical knowledge, pedagogical strategies, and ethical reasoning—thus emphasizing interdisciplinary thinking and reflective practice (Sperling et al., 2024). While these contributions have significantly advanced the conceptual landscape of AI literacy in education, much of the existing literature remains focused on operational skills. Relatively few studies explore the psychological or occupational dimensions of AI literacy, such as its impact on educators' motivation, emotional states, or overall professional experience.

Building on the reviewed literature, the present study adopts a comprehensive definition proposed by Laupichler et al. (2022), Ng et al. (2023), and Sperling et al. (2024). University faculty's AI literacy is defined as the systematic understanding and practical application of AI principles, tools, ethical considerations, and implementation strategies. Its core dimensions include technical proficiency, algorithmic thinking,



Note(s): TK= Technological Knowledge; CK= Content Knowledge; TPK= Technological Pedagogical Knowledge; TPAK= Technological Pedagogical Affective Knowledge; TPAC= Technological Pedagogical And Content Knowledge.

Source(s): Created by author

FIGURE 1
University faculty's use of AI.

interdisciplinary integration, sensitivity to educational equity, and awareness of AI's broader societal implications (Abulibdeh et al., 2024). Faculty members with higher AI literacy typically demonstrate more favorable attitudes toward technology and greater competence in integrating AI into their professional roles, which in turn enhances teaching quality and research productivity (Bewersdorff et al., 2025). Therefore, AI literacy should not be conceptualized solely as a technical skill set, but rather as a psychologically meaningful framework that captures university faculty's cognitive, emotional, and professional engagement with emerging technologies.

2.2 Self-determination theory

Self-Determination Theory (SDT), initially proposed by Ryan and Deci (2020), offers a foundational lens for examining human motivation and psychological well-being. At the core of SDT is the assertion that individuals possess three innate psychological needs—perceived autonomy, perceived competence, and perceived relatedness—whose fulfillment is essential for fostering intrinsic motivation, optimal functioning, and mental health (Janssen et al., 2013). Perceived autonomy involves volitional behavior and self-guided action; perceived competence reflects individuals' beliefs in their abilities to successfully perform tasks; and perceived relatedness pertains to the sense of meaningful connection and support from others (Deci and Ryan, 2000).

These needs have been consistently associated with key professional outcomes in educational settings, including job satisfaction, psychological well-being, and organizational commitment (Inigo and Raufaste, 2019). In higher education, faculty autonomy in selecting and utilizing digital tools has been shown to enhance intrinsic motivation (Zheng et al., 2025). Likewise, confidence in using educational technology can reduce anxiety and improve engagement in teaching activities (Klassen and Chiu, 2010), while the satisfaction of relatedness needs fosters emotional support and strengthens faculty members' sense of belonging within academic communities (Naidoo and Wagner, 2020).

Recent research has increasingly applied SDT to technology adoption contexts, particularly in exploring how faculty adapt to AI-integrated teaching and learning environments (Francis et al., 2024). These studies suggest that the degree to which AI innovations support or undermine psychological need satisfaction is essential in influencing educators' attitudes, behaviors, and well-being.

Taken together, SDT provides a comprehensive and empirically grounded lens for examining the psychological mechanisms that underlie faculty engagement with AI technologies. It also offers a compelling theoretical foundation for understanding how AI literacy—as both a cognitive and behavioral construct—can influence motivation, job satisfaction, and broader professional experiences.

2.3 Work–life balance

Work–life balance refers to an individual's ability to manage and reconcile the competing demands of professional and personal life in a satisfying and sustainable manner (Greenhaus and Beutell, 1985). University faculty often face significant role strain and time pressure due to their multiple responsibilities in teaching, research, and service (Pasamar et al., 2020). Effective time management and allocation of cognitive and emotional resources are thus critical to their psychological well-being (Pace et al., 2021).

AI technologies offer promising tools to alleviate academic workload and enhance efficiency, potentially affording faculty greater flexibility and control over their time (Bhojak et al., 2025). For example, automated grading systems and intelligent scheduling platforms can streamline repetitive tasks and improve task allocation, thereby helping faculty manage the boundary between work and personal life more effectively (Badri, 2024). However, in the absence of adequate institutional support, technological integration may give rise to new stressors—such as cognitive overload, digital fatigue, and increased anxiety—which may undermine rather than enhance work–life balance (Dorenkamp and Ruhle, 2019).

2.4 Job satisfaction

Job satisfaction is a multidimensional construct that reflects an individual's overall appraisal of their work experience, encompassing aspects such as task content, work environment, interpersonal relationships, and career development opportunities. It is widely recognized as a key indicator of professional well-being and engagement among faculty (Troesch and Bauer, 2017). Both extrinsic factors (e.g., salary, institutional policies) and intrinsic factors (e.g., teaching motivation, academic identity) influence job satisfaction levels (Layek and Koodamara, 2024).

In the context of increasing digitalization in higher education, AI literacy has emerged as an important predictor of job satisfaction (Jose et al., 2025). Faculty members with higher levels of AI literacy often report a greater sense of control, efficacy, and competence in navigating digital teaching environments, which in turn enhances engagement and fulfillment (Ji et al., 2025). Moreover, AI tools that automate routine tasks and streamline workflows can also boost productivity and reinforce a sense of accomplishment (Xia et al., 2022). However, disparities in technological readiness can lead to anxiety, information overload, and emotional exhaustion—factors that detract from overall job satisfaction (Li and Yu, 2022).

2.5 Technology acceptance

Technology acceptance is commonly defined as an individual's evaluation of and readiness to embrace new technologies, often framed by two key constructs: perceived usefulness and ease of use (Venkatesh and Davis, 2000). In educational settings, technology acceptance plays a pivotal role in shaping faculty members' adoption behavior and their emotional responses to digital tools (Scherer et al., 2019).

AI technologies, while potentially transformative, often present barriers to acceptance due to their complexity and ethical concerns related to privacy, transparency, and accountability (Bergdahl et al., 2023). Conversely, educators with high levels of technology acceptance are more likely to engage in active learning, integrate innovative tools into their teaching, and demonstrate greater openness to pedagogical experimentation (Racero et al., 2020).

Importantly, technology acceptance may also moderate the relationship between AI literacy and psychological outcomes. It can shape how faculty experience autonomy, competence, and emotional responses when implementing AI in their instructional and research practices (Antonietti et al., 2022). Understanding the role of technology acceptance is therefore essential for designing effective faculty development programs and for promoting sustainable and psychologically supportive AI integration in higher education.

3 Research model and hypothesis development

3.1 The impact of AI literacy on self-determination theory constructs: perceived autonomy, perceived competence, and perceived relatedness

AI literacy, beyond operational proficiency, represents faculty members' ability to self-direct technology use, critically evaluate AI-generated outputs, and integrate tools into pedagogical practices (Celik, 2023). High AI literacy equips teachers to select suitable AI functions, customize workflows, and adapt teaching strategies without relying heavily on external guidance (Chiu and Chai, 2020). This capability fosters perceived autonomy because decisions about technology use are internally regulated rather than externally imposed (Ryan and Deci, 2000). Autonomy emerges when individuals experience volition in aligning AI applications with their instructional goals, reducing feelings of technological constraint. Conversely, low AI literacy can lead to dependency on preset tools or institutional mandates, limiting choice and control. The presence of high AI literacy therefore strengthens the sense of ownership over instructional processes, enabling faculty to exercise freedom in technology adoption and thereby enhancing self-determined engagement in AI-enhanced education (Zawacki-Richter et al., 2019). Accordingly, we propose:

> *H1*: AI literacy has a significant positive effect on university faculty's perceived autonomy in using AI.

Perceived competence reflects the belief in one's ability to effectively perform tasks (Ryan and Deci, 2000). In AI-supported higher education, this belief is strengthened when teachers possess the necessary technical and cognitive skills to translate AI capabilities into academic outcomes (Xia et al., 2023). Faculty with high AI literacy can manage data-driven analytics, apply intelligent feedback, and integrate cross-platform resources efficiently. These abilities reduce uncertainty when facing complex tasks, reinforcing task mastery and professional efficacy (Wang et al., 2025). Mastery experiences with AI tools also create positive performance feedback loops, increasing confidence and willingness to undertake more challenging projects. In contrast, insufficient AI literacy may result in trial-and-error inefficiency, eroding competence perceptions. Thus, AI literacy operates as a foundational resource that transforms technical knowledge into tangible achievements, directly enhancing teachers' confidence in their technological and pedagogical capabilities. Therefore, we propose:

> *H2*: AI literacy has a significant positive effect on university faculty's perceived competence in using AI.

Perceived relatedness, as defined in Self-Determination Theory, reflects the need to feel connected to others and experience mutual support (Ryan and Deci, 2000). AI literacy enhances relatedness by enabling faculty to participate effectively in technology-mediated collaboration, such as co-teaching, shared resource creation, and engagement in virtual scholarly communities (Ng et al., 2023). Proficiency in AI tools facilitates smooth communication, efficient content sharing, and mutual problem-solving, which strengthen interpersonal trust and social bonds (Singh and Aziz, 2025). When teachers can competently navigate AI-enhanced platforms, they are more likely to contribute meaningfully to joint projects, receive peer recognition, and build sustained professional relationships. This socially embedded use of AI fosters organizational belonging and reinforces collective identity (Wang et al., 2025; Xia et al., 2022). Conversely, low AI literacy can hinder participation in collaborative environments, reducing opportunities for connection and support. Accordingly, we propose:

> *H3*: AI literacy has a significant positive effect on university faculty's perceived relatedness in using AI.

3.2 The impact of perceived autonomy on work–life balance and job satisfaction

Perceived autonomy reflects the extent to which individuals feel free to make choices and regulate their actions according to personal goals (Ryan and Deci, 2000). In academic contexts, this sense of volition enables faculty to organize tasks, prioritize responsibilities, and adjust teaching schedules in ways that align with personal circumstances (Van den Broeck et al., 2016). Within AI-assisted environments, autonomy manifests when teachers can independently determine how and when to integrate AI tools, select functionalities suited to their needs, and adapt outputs for specific pedagogical purposes (Wu et al., 2024). Such flexibility reduces time pressure, minimizes role conflict, and enhances the ability to allocate resources between work and personal life (Khawand and Zargar, 2022). Greater autonomy also supports proactive coping strategies, allowing educators to manage workload without compromising personal well-being (Fotiadis et al., 2019). These mechanisms explain why autonomy in AI use is likely to facilitate a more sustainable work–life balance for faculty members. Therefore, we propose:

> *H4*: University faculty's perceived autonomy in using AI has a significant positive effect on their work–life balance.

Autonomy in AI-supported teaching fosters a sense of control over instructional decisions, enabling faculty to select content, pedagogical strategies, and technological configurations that best serve their objectives (Gagné et al., 2022). This self-directed approach strengthens goal alignment, reinforces intrinsic motivation, and enhances professional purpose (Ma and Vu, 2024). In personalized teaching contexts, AI tools allow real-time adaptation of learning materials, automated assessment, and targeted feedback, enabling teachers to implement innovations without excessive external constraints (Cho and Jung, 2025). Positive experiences with such autonomy can increase satisfaction through improved student outcomes and professional recognition. By contrast, limited control over AI integration may generate frustration or reduce engagement. The ability to decide when and how to use AI therefore directly contributes to job satisfaction by reinforcing educators' sense of competence, self-worth, and alignment with institutional goals. Accordingly, we propose:

> *H5*: University faculty's perceived autonomy in using AI has a significant positive effect on their job satisfaction.

3.3 The impact of perceived competence on work–life balance and job satisfaction

Perceived competence describes individuals' confidence in their ability to meet situational demands effectively (Deci and Ryan, 2000). In AI-enhanced teaching environments, competent faculty can quickly identify appropriate tools, adapt them to diverse instructional needs, and transfer learned skills across tasks (Celik, 2023). This efficiency reduces trial-and-error inefficiencies, allowing more time for strategic planning and personal activities. Competence also supports better workload structuring, helping faculty manage the boundary between work and non-work domains (Yin and Huang, 2021). Automation capabilities, such as grading algorithms and scheduling assistants, further improve resource allocation and reduce repetitive tasks (Ayanwale et al., 2024). As educators gain mastery over AI, they are more likely to experience control over their professional routines, freeing time for personal commitments and improving overall life balance. Therefore, we propose:

> *H6*: University faculty's perceived competence in using AI has a significant positive effect on their work–life balance.

Competence contributes to higher self-efficacy, enabling educators to set challenging goals and maintain confidence in achieving them (Klassen and Chiu, 2010). In AI-supported contexts, skill growth can lead to improved teaching quality, innovative research outputs, and enhanced operational control (Wang et al., 2025). Mastery experiences with AI reduce apprehension toward technological change, facilitating adaptability and reducing work-related stress (Hofer et al., 2021). Competent educators are also more likely to receive positive feedback from peers, students, and institutions, reinforcing a sense of accomplishment. Personalized AI applications, such as adaptive learning platforms, can further increase recognition by highlighting the impact of teachers' expertise on student performance (Molefi et al., 2024). These factors collectively strengthen job satisfaction by fulfilling psychological needs for achievement and professional growth. Accordingly, we propose:

> *H7*: University faculty's perceived competence in using AI has a significant positive effect on their Job Satisfaction.

3.4 The impact of perceived relatedness on work–life balance and job satisfaction

Within Self-Determination Theory, relatedness reflects the need to feel connected and supported by others (Deci and Ryan, 2000). In AI-integrated higher education, strong relatedness allows faculty to form cooperative networks that buffer the uncertainty and stress of adopting new technologies (Prenger et al., 2017). Engagement in collaborative teaching, co-development of AI-supported learning resources, and participation in online academic communities facilitates the exchange of both technical and emotional resources (Ng et al., 2023). Such networks reduce the cognitive load of technology use, enhance mutual trust, and create a supportive environment for managing workload. Social bonds also aid emotional regulation and promote adaptive coping, which preserves energy for non-work activities (Ertiö et al., 2024). Through these mechanisms, perceived relatedness enables faculty to sustain an equilibrium between professional and personal life, thereby enhancing work–life balance. Therefore, we propose:

> *H8*: University faculty's perceived relatedness in using AI has a significant positive effect on their work–life balance.

Supportive professional relationships contribute to a sense of belonging and recognition, which strengthens academic identity and motivation (Klassen et al., 2012). In AI-supported environments, shared platforms enable faculty to collaborate on innovative teaching designs, exchange strategies for AI integration, and celebrate collective achievements (Dilek et al., 2025). These interactions reinforce shared goals, reduce isolation, and increase emotional engagement in academic work (Wang et al., 2025). A strong sense of relatedness also provides a stable psychological foundation for navigating challenges in AI adoption, as mutual support reduces stress and enhances confidence. Furthermore, expanded academic networks foster opportunities for interdisciplinary collaboration and professional growth, which in turn heightens organizational belonging (Singh and Aziz, 2025). By reinforcing both emotional satisfaction and professional accomplishment, perceived relatedness serves as a psychological driver of job satisfaction in AI-integrated teaching and research contexts. Therefore, we propose:

> *H9*: University faculty's perceived relatedness in using AI has a significant positive effect on their job satisfaction.

3.5 The impact of work–life balance on job satisfaction

Work–life balance is essential for sustaining professional well-being because it mitigates role conflict and reduces burnout (Prasad and Pasupathi, 2025). Faculty who manage work and personal responsibilities effectively can maintain emotional stability and preserve cognitive resources for teaching and research (Cao et al., 2020). Balanced schedules promote a sense of control, enabling educators to respond more constructively to academic challenges. In AI-integrated environments, automation tools streamline grading, scheduling, and data analysis, thereby reducing time spent on repetitive tasks and freeing capacity for personal and family activities (Meharunisa et al., 2024). This efficiency creates a reinforcing cycle in which improved balance enhances overall satisfaction with work, and higher satisfaction further motivates effective time management. Faculty who sustain this equilibrium often report greater role clarity, stronger engagement, and a deeper sense of professional accomplishment (Wei and Ye, 2022). Therefore, we propose:

> *H10*: University faculty's job satisfaction is significantly and positively influenced by their work–life balance.

3.6 The moderating role of technology acceptance

Technology acceptance, as conceptualized in the Technology Acceptance Model, captures individuals' perceptions of ease of use, perceived usefulness, and readiness to engage with new tools (Venkatesh et al., 2003). In the context of AI-enhanced higher education, high acceptance enables faculty to translate AI literacy into meaningful, autonomous use of technology. When teachers perceive AI as valuable and manageable, they are more likely to experiment with functions, adapt tools to personal teaching styles, and initiate independent problem-solving (Chiu et al., 2024). This strengthens their perceived control over technological processes and their freedom to design instructional approaches. Conversely, low acceptance can lead to avoidance behaviors, heightened anxiety, and underutilization of existing AI skills (Teo, 2011). In such cases, AI literacy may remain a latent capability rather than an active driver of autonomy. Technology acceptance thus shapes the psychological translation of AI knowledge into self-determined teaching behaviors (Dahri et al., 2024). Therefore, we propose:

> *H11*: Technology acceptance moderates the positive relationship between AI literacy and perceived autonomy in using AI.

Faculty with high technology acceptance are more inclined to devote time and cognitive effort to mastering AI systems, which allows them to achieve operational proficiency and receive timely performance feedback (Scherer et al., 2019). This active engagement fosters a deeper understanding of AI functionalities and facilitates effective application in both teaching and research. When acceptance is high, AI literacy is more efficiently converted into perceived competence, reinforcing professional self-efficacy and enabling teachers to tackle complex academic tasks with confidence (Wang et al., 2025). In contrast, low acceptance often leads to neglect of learning opportunities, reluctance to apply existing AI knowledge, and limited skill growth (Sumak et al., 2011). Even when technical capabilities exist, resistance toward technology may hinder the translation of knowledge into effective practice. Thus, technology acceptance determines the extent to which AI literacy can be transformed into a tangible sense of competence in professional contexts. Therefore, we propose:

> *H12*: Technology acceptance moderates the positive relationship between AI literacy and perceived competence in using AI.

Perceived relatedness in AI integration depends not only on technical knowledge but also on willingness to engage in technology-mediated collaboration (Ng et al., 2023). Faculty with high technology acceptance tend to view AI platforms as effective tools for interaction, resource sharing, and collective problem-solving (Kaliisa et al., 2022). This positive orientation encourages participation in interdisciplinary networks, joint curriculum design, and virtual academic communities, which strengthens interpersonal bonds and mutual trust. In contrast, low acceptance can result in avoidance of AI-based interactions, reduced collaborative initiatives, and diminished exposure to diverse perspectives (Teo and Noyes, 2014). Over time, such withdrawal limits opportunities for social support and weakens organizational connectedness. By shaping the frequency and quality of AI-mediated exchanges, technology acceptance influences how AI literacy contributes to the satisfaction of relatedness needs and the development of a sense of belonging in academic settings. Therefore, we propose:

> *H13*: Technology acceptance moderates the positive relationship between AI literacy and perceived relatedness in using AI.

These hypothesized relationships are synthesized in the research model depicted in Figure 2.

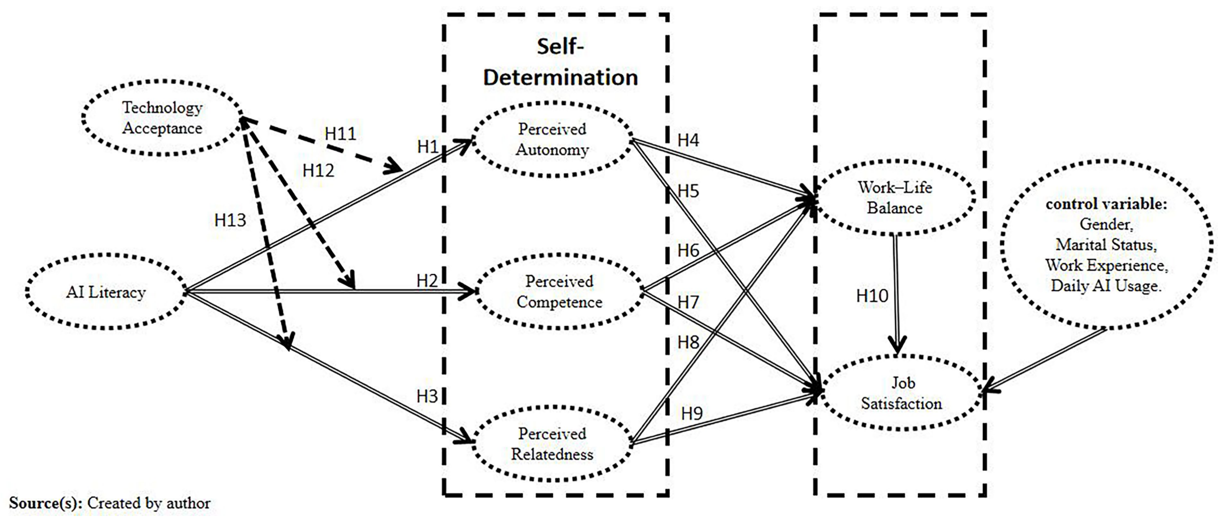


FIGURE 2
Research model.

4 Method

4.1 Participants

Data for this study were collected via an online questionnaire administered through Wenjuanxing[1], a commonly used platform in China. The instrument employed a seven-point Likert scale and was distributed using a stratified random sampling strategy. Participants were recruited through the platform's representative sampling service, ensuring diverse and demographically balanced responses. All surveys were completed electronically. To incentivize participation, respondents who submitted valid questionnaires received a monetary reward of 5 RMB. To ensure data quality and participant relevance, the study implemented multiple inclusion and exclusion criteria. (1) Screening Questions: At the beginning of the questionnaire, two mandatory screening items were included to confirm eligibility: (a) "What is your current occupation?"—only respondents selecting "university faculty member" were included; and (b) "Do you have basic experience using AI tools?"—only those responding "Yes" proceeded to the main section. Failure to meet either criterion led to immediate exclusion. (2) Attention Check: A directed-response item was embedded mid-questionnaire (e.g., "To confirm you are paying attention, please select 'Strongly disagree' for this item."). Participants who failed to respond as instructed were excluded from the final dataset. (3) Responses exhibiting identical selections across all Likert-scale items were flagged for satisficing behavior or inattentive responding and subsequently removed from the analysis. In total, 543 questionnaires were received. After removing 32 invalid responses based on the criteria above, 511 valid responses remained, resulting in an effective response rate of 94.11%. A summary of participant demographic characteristics is presented in Table 2.

1 https://www.wjx.cn/vm/eAGu0vk.aspx

TABLE 2 Demographic characteristics of participants.

Variable	Category	Frequency	Percentage
Gender	Male	209	40.90%
	Female	302	59.10%
Marital status	Married	288	56.36%
	Unmarried	223	43.64%
Work experience	0–5 years	110	21.53%
	6–10 years	198	38.75%
	11–15 years	130	25.44%
	16–20 years	47	9.20%
	More than 21 years	26	5.09%
Daily AI usage	0–1 h	178	34.83%
	1–3 h	143	27.98%
	3–5 h	115	22.50%
	5–7 h	53	10.37%
	More than 7 h	22	4.31%

Source(s): Created by author.

4.2 Measures

Measurement instruments in this research were derived from established scales with prior empirical validation across domestic and international studies (Haw et al., 2024; Lin et al., 2025; Mulyani et al., 2021; Wang et al., 2023; Yildirim et al., 2024). All scales were appropriately modified to align with the specific context of university faculty. The core variables measured in this study included AI literacy (AIL), perceived autonomy (PA), perceived competence (PC), perceived relatedness (PR), work–life balance (WLB), job satisfaction (JS), and technology acceptance (TA). Detailed information regarding the measurement dimensions, sample items, and sources of each scale is provided in Appendix 1.

4.3 Common method bias assessment

Given that all variables in this study were measured through cross-sectional self-reported questionnaires, procedural remedies were implemented to reduce the likelihood of common method variance (CMV). These included ensuring respondent anonymity, counterbalancing the order of measurement items, and embedding an attention-check question to minimize socially desirable responding and inattentive answering. To statistically assess CMV, Harman's single-factor test was performed by entering all measurement items into an unrotated exploratory factor analysis. Results showed that the first factor explained 38.844% of the total variance, which is below the commonly accepted threshold of 40%, suggesting that CMV was not a severe issue (Podsakoff et al., 2003). In addition, a common latent factor approach was employed in SmartPLS to detect potential method effects. The variance inflation factors (VIFs) for all items were below 3.3, indicating that multicollinearity was not a concern and CMV was unlikely to bias the study's results.

5 Data analysis and results

This study constructed a structural equation model (SEM) to examine the pathways through which AI literacy influences work–life balance and job satisfaction among university faculty. Descriptive statistics and preliminary analyses were conducted using SPSS 27.0, followed by SEM and path analysis using SmartPLS 4.0. In SEM methodology, covariance-based SEM (CB-SEM) is commonly used for confirmatory theory testing and evaluating global model fit, whereas partial least squares SEM (PLS-SEM) focuses on maximizing predictive accuracy and is especially appropriate for complex models, non-normal data, and exploratory research contexts. In this study, PLS-SEM was selected for four primary reasons: (1) the model includes seven latent constructs, multiple indicators, and a moderating effect, resulting in high structural complexity; (2) preliminary diagnostics revealed slight deviations from multivariate normality; (3) the research adopts an exploratory path-testing orientation aimed at extending rather than strictly confirming existing theory; and (4) PLS-SEM offers robustness and predictive power particularly suited to emerging research domains such as AI literacy in higher education. This choice is consistent with established methodological guidelines (Dash and Paul, 2021), which recommend PLS-SEM in early-stage theoretical model development and when prediction is a key goal.

5.1 Measurement model assessment

Validity was examined through three dimensions: content, convergent, and discriminant validity. Content validity was ensured by adapting measurement items from extensively validated scales in prior research. Convergent validity was supported by factor loadings (≥0.70) (Raykov and Marcoulides, 2019) and Average Variance Extracted ($AVE \geq 0.50$), in line with Cheung et al. (2024). Discriminant validity was evaluated using the Heterotrait–Monotrait ratio (HTMT), with the recommended threshold set below 0.85.

As shown in Tables 3, 4, the measurement model demonstrated satisfactory reliability and validity. All latent variables exhibited Cronbach's α values ranging from 0.903 to 0.947 and CR values ranging from 0.928 to 0.955, indicating strong internal consistency. Factor loadings fell within the range of 0.777 to 0.879, while AVE values ranged between 0.664 and 0.732, satisfying the criteria for convergent validity. HTMT values for all variable pairs were below 0.85, indicating good discriminant validity among the constructs.

5.2 Structural model analysis

Before testing the hypothesized structural relationships, four control variables—gender, marital status, work experience, and daily AI usage—were included in the model to account for potential confounding effects on job satisfaction. The results showed that none of these control variables had a statistically significant impact on job satisfaction (all $p > 0.05$), indicating that the subsequent path estimates are unlikely to be biased by these demographic or usage-related factors.

5.2.1 Model explanatory power and predictive relevance

Table 5 presents the coefficients of determination (R^2) and predictive relevance (Q^2) for the endogenous variables. The R^2 values for perceived autonomy (0.315), perceived competence (0.337), perceived relatedness (0.345), work–life balance (0.435), and job satisfaction (0.253) all reached acceptable levels, with the explanatory power for work–life balance being particularly strong. All Q^2 values were greater than zero (ranging from 0.218 to 0.342), indicating that the model demonstrates satisfactory predictive relevance. These results support the robustness and theoretical validity of the model.

5.2.2 Path coefficient analysis and hypothesis testing

Figure 3 illustrates the structural model outcomes, indicating that AI literacy significantly and positively influenced perceived autonomy ($\beta = 0.404, p < 0.001$), perceived competence ($\beta = 0.468, p < 0.001$), and perceived relatedness ($\beta = 0.432, p < 0.001$), providing strong support for H1, H2, and H3. These results suggest that higher levels of AI literacy among university faculty are associated with enhanced experiences of self-control, competence fulfillment, and social connectedness.

Further analysis revealed that Perceived Autonomy ($\beta = 0.278, p < 0.001$), Competence ($\beta = 0.259, p < 0.001$), and Relatedness ($\beta = 0.251, p < 0.001$) all significantly and positively influenced work–life balance, supporting H4, H6, and H8. These findings indicate that when faculty members' psychological needs are met, they are more capable of balancing teaching, research, and personal life, which helps reduce role conflicts. Regarding the impact on job satisfaction, only perceived autonomy showed a significant positive effect ($\beta = 0.221, p < 0.001$), supporting H5. The effects of perceived competence ($\beta = 0.081, p = 0.186$) and perceived relatedness ($\beta = 0.042, p = 0.471$) were not statistically significant, and thus H7 and H9 were not supported. Several potential explanations for these results have been suggested in the literature. Li et al. (2025) noted that university faculty's job satisfaction is strongly influenced by organizational factors such as performance evaluations, work environment, promotion pathways, and career support, which may dilute the direct impact of perceived competence. Chang et al. (2024) emphasized that the pressures brought by technological penetration may counteract its positive effects. Lyu and Zhu (2019), as well as Yang and Ling (2023), argued that university teaching tends to be relatively independent, with lower frequencies of social interaction, making it difficult for perceived relatedness to have a significant effect on satisfaction. Kim (2024) further pointed out that current AI tools in education primarily focus on enhancing individual efficiency, while their capacity to support social interaction and collaborative work remains underdeveloped.

Additionally, work–life balance was found to be a significant positive predictor of job satisfaction ($\beta = 0.255, p < 0.001$), providing support for H10. This suggests that faculty members' positive evaluations of their work are closely tied to their ability to effectively integrate work and life roles.

5.2.3 Moderating effect analysis

As shown in Figure 3, technology acceptance significantly moderated the relationships between AI literacy and perceived autonomy ($\beta = 0.116, p = 0.007$), perceived competence ($\beta = 0.114, p = 0.004$), and perceived relatedness ($\beta = 0.199, p < 0.001$). These results provide empirical support for H11, H12, and H13. The findings suggest that higher levels of Technology Acceptance enhance the positive psychological impact of AI literacy. In other words, when teachers are more willing to embrace AI technologies, their AI literacy is more effectively translated into positive perceptions of autonomy, competence, and social connectedness.

To further examine and visually illustrate the moderating role of technology acceptance, this study plotted interaction diagrams (see Figure 4) depicting the relationships between AI literacy and perceived autonomy, competence, and relatedness at different levels of technology acceptance (low = −1 SD; high = +1 SD). As shown in Figure 4, a significant moderation effect was observed, wherein the positive pathways from AI literacy to basic psychological need satisfaction were amplified at higher levels of technology acceptance.

The results presented in Table 6 indicate that the effects of AI literacy on perceived autonomy ($\beta = 0.292, p < 0.001$), perceived competence ($\beta = 0.357, p < 0.001$), and perceived relatedness ($\beta = 0.226, p < 0.001$) are significant at low levels of technology acceptance. At moderate levels of technology acceptance, these effects become stronger for perceived autonomy ($\beta = 0.410, p < 0.001$), perceived competence ($\beta = 0.470, p < 0.001$), and perceived relatedness ($\beta = 0.418, p < 0.001$). At high levels of technology acceptance, the positive influence of AI literacy reaches its peak—perceived autonomy ($\beta = 0.528, p < 0.001$), perceived competence ($\beta = 0.583, p < 0.001$), and perceived relatedness ($\beta = 0.610, p < 0.001$). These findings suggest that the positive impact of AI literacy on individuals' basic psychological needs becomes progressively stronger

TABLE 3 Results of reliability and convergent validity testing.

Latent variable	Measurement items	Mean	Standard deviation	Factor loading	Cronbach's α	CR	AVE
AIL	AIL1	4.885	1.530	0.837	0.947	0.955	0.704
	AIL2	4.867	1.573	0.848			
	AIL3	4.793	1.583	0.840			
	AIL4	4.824	1.509	0.864			
	AIL5	4.765	1.548	0.806			
	AIL6	4.863	1.486	0.841			
	AIL7	4.691	1.529	0.826			
	AIL8	4.806	1.597	0.858			
	AIL9	4.828	1.510	0.829			
PA	PA1	4.941	1.544	0.872	0.906	0.930	0.726
	PA2	4.886	1.613	0.879			
	PA3	4.810	1.584	0.848			
	PA4	4.869	1.483	0.848			
	PA5	4.773	1.499	0.813			
PC	PC1	4.806	1.483	0.855	0.908	0.932	0.732
	PC2	4.750	1.458	0.861			
	PC3	4.943	1.567	0.847			
	PC4	4.932	1.595	0.858			
	PC5	4.693	1.502	0.856			
PR	PR1	4.722	1.474	0.856	0.903	0.928	0.719
	PR2	4.624	1.501	0.850			
	PR3	4.708	1.432	0.831			
	PR4	4.722	1.473	0.856			
	PR5	4.841	1.507	0.848			
WLB	WLB1	4.730	1.569	0.841	0.941	0.951	0.709
	WLB2	4.793	1.488	0.832			
	WLB3	4.605	1.545	0.849			
	WLB4	4.712	1.523	0.855			
	WLB5	4.683	1.544	0.856			
	WLB6	4.769	1.533	0.855			
	WLB7	4.765	1.525	0.841			
	WLB8	4.724	1.505	0.804			
JS	JS1	5.121	1.392	0.839	0.903	0.928	0.721
	JS2	5.155	1.435	0.870			
	JS3	5.182	1.389	0.850			
	JS4	5.147	1.382	0.850			
	JS5	5.213	1.464	0.838			
TA	TA1	4.781	1.480	0.830	0.937	0.947	0.664
	TA2	4.777	1.494	0.806			
	TA3	4.886	1.397	0.833			
	TA4	4.810	1.464	0.837			
	TA5	4.863	1.512	0.840			
	TA6	4.977	1.389	0.818			
	TA7	4.853	1.480	0.809			
	TA8	4.820	1.447	0.781			
	TA9	4.814	1.443	0.777			

Source(s): Created by author.

TABLE 4 Results of discriminant validity testing.

Variable	AIL	PA	PC	PR	WLB	JS	TA
AIL							
PA	0.542						
PC	0.588	0.627					
PR	0.565	0.573	0.608				
WLB	0.564	0.600	0.600	0.580			
JS	0.389	0.475	0.410	0.373	0.480		
TA	0.406	0.430	0.393	0.403	0.446	0.627	

Source(s): Created by author.

TABLE 5 Explanatory power (R^2) and predictive relevance (Q^2) of the structural model.

Endogenous Latent Variables	R^2	Q^2
PA	0.315	0.303
PC	0.337	0.323
PR	0.345	0.327
WLB	0.435	0.342
JS	0.253	0.218

Source(s): Created by author.

as technology acceptance increases, providing further support for hypotheses H11, H12, and H13.

6 Discussion

Grounded in Self-Determination Theory (SDT), this study developed a structural model to investigate how AI literacy influences university faculty's work–life balance and job satisfaction by satisfying three fundamental psychological needs: perceived autonomy, perceived competence, and perceived relatedness. Additionally, this study examined the moderating role of technology acceptance. The empirical results largely supported the proposed hypotheses, demonstrating that AI literacy is not merely a technical skill but a critical resource for activating intrinsic motivation and enhancing psychological well-being among teachers.

The findings revealed that AI literacy significantly and positively affects perceived autonomy, perceived competence, and perceived relatedness (H1–H3 supported), validating the pathway of "skill enhancement → psychological need satisfaction → motivational activation." These results are consistent with prior studies (Ji et al., 2025; Pauw et al., 2022; Pelau et al., 2021), which emphasized that AI literacy not only improves technological performance but also strengthens teachers' sense of instructional control, competence recognition, and social connectedness, ultimately enhancing their professional engagement.

Among the three psychological needs, perceived autonomy emerged as a significant positive predictor of both work–life balance and job satisfaction (H4 and H5 supported), corroborating Rahimi et al. (2024), who highlighted the pivotal role of autonomy in high-autonomy, high-demand professions. When teachers have greater decision-making power and scheduling flexibility in their use of AI technologies, they are better able to manage task pacing and balance multiple roles, thereby improving their job satisfaction (Deci et al., 2017).

In contrast, although perceived competence and perceived relatedness significantly enhanced work–life balance (H6 and H8 supported), their direct effects on job satisfaction were not significant (H7 and H9 not supported). One possible explanation lies in the evolving institutional and technological context of higher education. Many university performance evaluation systems still place greater emphasis on research output, grant acquisition, and autonomy in teaching innovation, while offering limited recognition for competence gains derived from technological adaptation (Singh et al., 2022). This focus may weaken the intrinsic satisfaction associated with improved competence, especially when such competence is perceived as an instrumental, externally driven requirement rather than a source of long-term professional pride (Lai and Jin, 2021; Wang et al., 2022). In AI-integrated teaching environments, competence increasingly reflects rapid adaptation and operational efficiency, characteristics that are transient and dependent on continuous technological updates, thereby reducing their potential to sustain job satisfaction (Zhou J. S. et al., 2024; Zhou T. et al., 2024). Similarly, the non-significant direct effect of perceived relatedness on job satisfaction may be linked to the functional limitations of current AI tools in fostering meaningful social connections. Although digital platforms facilitate communication, they tend to prioritize efficiency and task completion over relational depth (Yu, 2024; Zagni et al., 2025). As face-to-face collaboration is increasingly replaced by asynchronous or AI-mediated exchanges, opportunities for spontaneous peer support, emotional bonding, and informal knowledge sharing—critical components of professional fulfillment—are reduced (Turk et al., 2022). Furthermore, many AI tools in educational settings lack embedded collaborative features designed to build interpersonal trust and mutual support, thereby limiting their capacity to enhance relatedness in ways that translate directly into job satisfaction (Pelau et al., 2021). Taken together, these findings suggest that in technology-rich educational environments, competence and relatedness may contribute to job satisfaction mainly through indirect pathways—most notably by improving work–life balance—rather than through direct influence. This underscores the need for institutional reward systems that explicitly acknowledge competence development in technology adoption and for AI tools that integrate richer collaborative functions to strengthen professional relatedness (Rahimi et al., 2024; Singh et al., 2022).

Importantly, work–life balance significantly predicted job satisfaction (H10 supported), aligning with the findings of Landolfi et al. (2021), who emphasized the crucial role of life balance in constructing professional well-being. University faculty members who can effectively integrate their work and life roles are more likely to experience emotional stability and greater happiness. This result aligns with the work–family conflict framework outlined by Greenhaus and Beutell (1985), which posits that effective life balance can mitigate job stress and enhance satisfaction. Although perceived competence and perceived relatedness did not directly predict job satisfaction, they still exerted indirect effects through their positive contributions to work–life balance.

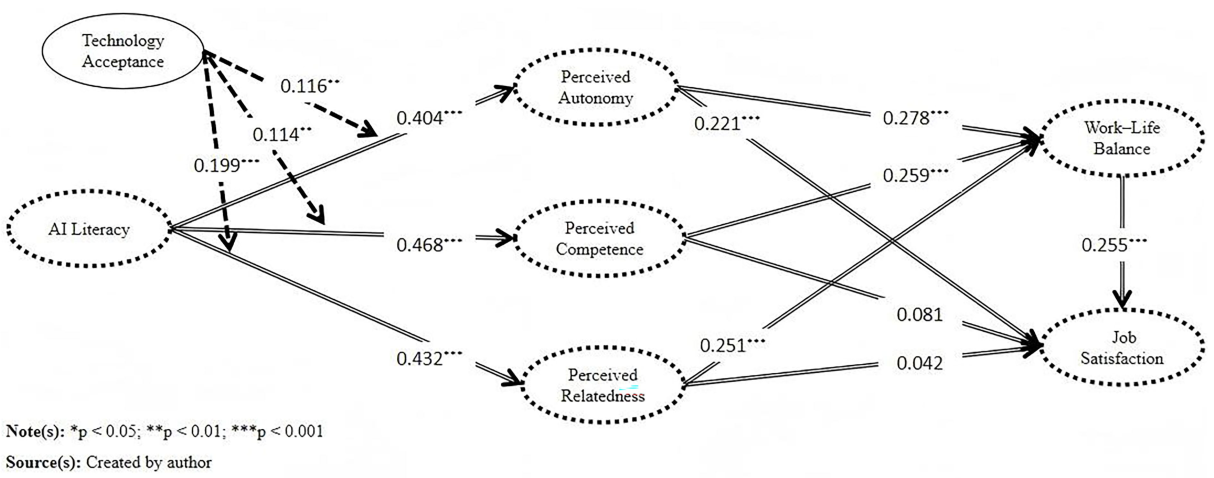


FIGURE 3
Path analysis results of the structural equation model.

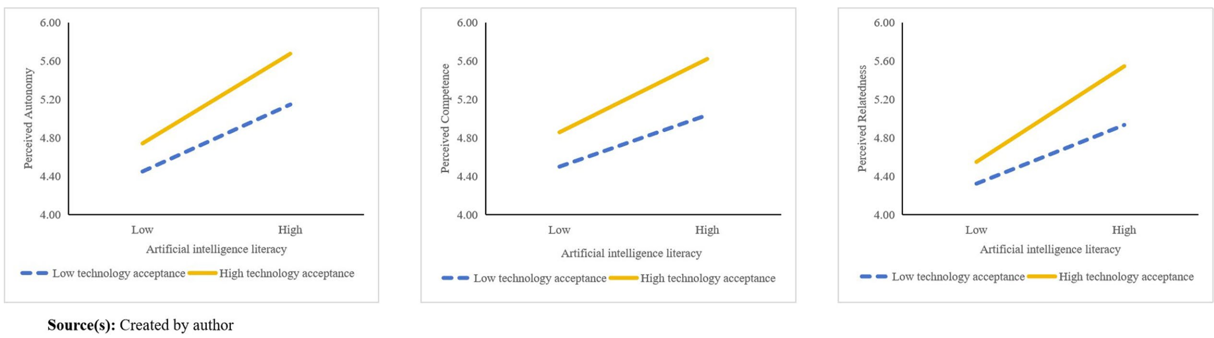


FIGURE 4
The moderating effect of technology acceptance.

TABLE 6 Moderating effect at different levels.

Independent variable	Dependent variable	TA	β	95%CI		p-value
AIL	PA	Low	0.292	0.176	0.407	***
		Medium	0.410	0.330	0.490	***
		High	0.528	0.421	0.635	***
	PC	Low	0.357	0.244	0.469	***
		Medium	0.470	0.392	0.547	***
		High	0.583	0.478	0.687	***
	PR	Low	0.226	0.118	0.333	***
		Medium	0.418	0.344	0.492	***
		High	0.610	0.510	0.710	***

*p < 0.05, **p < 0.01, ***p < 0.001.
Source(s): Created by author.

Furthermore, technology acceptance significantly moderated the relationships between AI literacy and the three basic psychological needs (H11–H13 supported). This finding is consistent with the "cognition–attitude–behavior" sequence in the technology acceptance Model (Belletier et al., 2018), suggesting that technology acceptance amplifies the positive psychological effects of AI literacy. Without sufficient confidence in and identification with AI tools, even teachers with high technical competence may struggle to fully activate psychological resources. Therefore, improving teachers' AI literacy should be accompanied by efforts to enhance their technology acceptance, including providing scenario-based training, cognitive empowerment, and emotional support to strengthen their recognition of and willingness to engage with AI tools.

7 Contributions, limitations, and future research directions

7.1 Theoretical contributions

Grounded in Self-Determination Theory (SDT), the present research constructed a structural framework to examine how AI literacy shapes university faculty's work–life balance and job satisfaction by fulfilling three fundamental psychological needs: perceived autonomy, perceived competence, and perceived relatedness. Additionally, the model integrated the technology acceptance as a moderating factor to systematically account for the psychological processes linking AI literacy to faculty well-being (Venkatesh et al., 2003). The theoretical contributions of this work are summarized in three key aspects:

First, this study extends the psychological conceptualization of AI literacy. Previous research has primarily regarded AI Literacy as an external technical competence or a performance indicator in educational settings (Celik, 2023; Chiu and Chai, 2020). In contrast, this study redefines AI Literacy as a psychological resource that activates intrinsic motivation. It demonstrates that AI Literacy can enhance work–life balance and job satisfaction by satisfying the psychological needs for autonomy, competence, and relatedness. The moderating effect of technology acceptance further reveals the critical role of cognitive attitudes in transforming technical literacy into motivational resources (Dahri et al., 2024), thereby expanding the theoretical boundaries and application pathways of AI literacy.

Second, this study reconstructs the pathway mechanisms and functional boundaries of SDT in high-technology environments. The findings indicate heterogeneous effects of the three psychological needs on faculty well-being: Perceived autonomy directly influences work–life balance and job satisfaction, while perceived competence and perceived relatedness primarily exert indirect effects through work–life balance (Ng et al., 2023; Van den Broeck et al., 2016). Furthermore, in AI-intensive teaching contexts, the direct impacts of competence and relatedness on job satisfaction were found to be non-significant, possibly due to the instrumental evaluation and function-oriented socialization patterns introduced by AI tools (Yu, 2024; Zagni et al., 2025). These results not only validate the context-dependency of SDT but also provide new insights for adapting and extending the theory in digital educational environments.

Third, this study advances the theoretical integration between educational technology and motivational psychology. The proposed integrated model—AI literacy → technology acceptance → psychological needs → job satisfaction—bridges the cognition–attitude mechanism of the technology acceptance Model with the motivation–well-being mechanism of SDT (Deci and Ryan, 2000; Venkatesh et al., 2003), offering a comprehensive framework to explain the interaction between individual behavior and psychological states in technology-driven settings.

7.2 Practical implications

This study offers the following practical recommendations for developing university faculty competencies and promoting the effective integration of educational technologies:

First, enhancing AI literacy should address both technical capabilities and psychological adaptability. Teacher training programs should not only develop operational skills and task optimization abilities but also cultivate the capacity for flexible technology transfer across varied instructional contexts, while fostering motivation and adaptability to change. Problem-based, project-oriented instructional activities can provide authentic problem-solving experiences, enabling teachers to build confidence and develop a sustained willingness to adopt AI tools.

Second, interventions should specifically target the indirect effects of competence and relatedness. Universities can integrate AI training with clear career development pathways, enabling teachers to enhance their skills while aligning them with professional growth trajectories, thereby strengthening competence satisfaction. At the same time, the design of collaborative AI-enabled teaching tools can promote cross-disciplinary resource sharing and foster the development of virtual academic communities, encouraging experience exchange and emotional connections that enhance relatedness and organizational belonging.

Third, it is essential to strengthen teachers' technology acceptance to facilitate the internalization of AI literacy as a motivational driver. Role modeling, case-based learning, and experiential training can reduce uncertainty and enhance acceptance. Establishing "AI Empowerment Facilitator" roles or teacher learning communities can leverage positive peer influence to encourage proactive use of AI in teaching and research.

Fourth, teacher support policies should evolve toward personalized and continuous interventions. Recognizing differences in psychological responses across academic ranks, disciplines, and age groups, universities should implement stratified and targeted support systems. For example, younger faculty may benefit from growth-oriented feedback and belongingness support, while senior faculty may require greater flexibility, recognition, and opportunities for legacy building. Such tailored approaches can better align AI literacy development with career advancement goals.

7.3 Research limitations

The primary data source for this study was self-reported questionnaires completed by university faculty in Mainland China, which may introduce potential biases such as self-report bias and socially desirable responding. Although common method bias was tested using Harman's single-factor method and found to be non-significant, the cross-sectional design limits the ability to capture dynamic changes and infer causal relationships over time. In addition, the current model did not control for or differentiate key job-related and demographic variables, such as teaching workload, academic rank, and disciplinary background, which may confound the relationship between AI literacy and job satisfaction. The absence of such controls limits the precision of the path estimates. Moreover, other unmeasured contextual variables (e.g., organizational support, institutional fairness, leadership styles, technology infrastructure) and personal characteristics (e.g., prior AI experience) may influence psychological need satisfaction and work-related outcomes, potentially affecting the model's external validity.

7.4 Directions for future research

Future research should address these limitations in several ways. First, incorporating multi-source data—such as teaching logs, platform usage records, AI interaction trajectories, and classroom observations—would improve measurement validity and reduce the reliance on self-reported data. Second, adopting longitudinal or intervention designs would allow researchers to track the dynamic evolution of AI literacy and assess its long-term effects on teachers' psychological states and professional well-being, thereby strengthening causal inference. Third, integrating multi-level contextual factors (e.g., organizational support, institutional fairness, leadership styles, technology infrastructure) and individual difference variables (e.g., academic rank, disciplinary background, prior AI experience) as covariates in the structural model would refine the precision of the estimates and illuminate potential moderating mechanisms. Finally, subgroup analyses could identify heterogeneous pathways in psychological need satisfaction across different teacher groups, offering tailored evidence for faculty development policies and technology-driven interventions.

Data availability statement

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

Author contributions

LH: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. YZ: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This work was supported by the Yunnan Provincial Undergraduate Colleges and Universities Education and Teaching Reform Research Project (Grant no. JG2025015) and the Humanities and Social Sciences Project of the Chongqing Municipal Education Commission (Grant no. 24SKSZ073).

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.

Generative AI statement

The authors declare that no Gen AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2025.1669247/full#supplementary-material

References

Abulibdeh, A., Zaidan, E. S., and Abulibdeh, R. (2024). Navigating the confluence of artificial intelligence and education for sustainable development in the era of industry 4.0: challenges, opportunities, and ethical dimensions. *J. Clean. Prod.* 437:140527. doi: 10.1016/j.jclepro.2023.140527

Al-Abdullatif, A. M. (2024). Modeling teachers' acceptance of generative artificial intelligence use in higher education: the role of AI literacy, intelligent TPACK, and perceived trust. *Educ. Sci.* 14:1209. doi: 10.3390/educsci14111209

Antonietti, C., Cattaneo, A., and Amenduni, F. (2022). Can teachers' digital competence influence technology acceptance in vocational education? *Comput. Human Behav.* 132:107266. doi: 10.1016/j.chb.2022.107266

Ayanwale, M. A., Adelana, O. P., Molefi, R. R., Adeeko, O., and Ishola, A. M. (2024). Examining artificial intelligence literacy among pre-service teachers for future classrooms. *Comput. Educ. Open* 6:100179. doi: 10.1016/j.caeo.2024.100179

Badri, S. K. Z. (2024). Evaluating higher education (HE) work design to sustain academics' work-life balance. *J. Appl. Res. High. Educ.* 16, 134–149. doi: 10.1108/jarhe-04-2022-0116

Belletier, C., Robert, A., Moták, L., and Izaute, M. (2018). Toward explicit measures of intention to predict information system use: an exploratory study of the role of implicit attitudes. *Comput. Human Behav.* 86, 61–68. doi: 10.1016/j.chb.2018.04.029

Bergdahl, J., Latikka, R., Celuch, M., Savolainen, I., Mantere, E. S., Savela, N., et al. (2023). Self-determination and attitudes toward artificial intelligence: cross-national and longitudinal perspectives. *Telemat. Inform.* 82:102013. doi: 10.1016/j.tele.2023.102013

Bewersdorff, A., Hornberger, M., Nerdel, C., and Schiff, D. S. (2025). AI advocates and cautious critics: how AI attitudes, AI interest, use of AI, and AI literacy build university students' AI self-efficacy. *Comput. Educ.* 8:100340. doi: 10.1016/j.caeai.2024.100340

Bhojak, N. P., Momin, M., Jani, D., and Mathur, A. (2025). Enhancing teachers' job satisfaction through the artificial intelligence utilization. *J. Appl. Res. High. Educ.* doi: 10.1108/jarhe-03-2024-0126

Cao, C., Shang, L., and Meng, Q. (2020). Applying the job demands-resources model to exploring predictors of innovative teaching among university teachers. *Teach. Teach. Educ.* 89:103009. doi: 10.1016/j.tate.2019.103009

Celik, I. (2023). Exploring the determinants of artificial intelligence (AI) literacy: digital divide, computational thinking, cognitive absorption. *Telematics Inform.* 83:102026. doi: 10.1016/j.tele.2023.102026

Chang, P. C., Zhang, W. H., Cai, Q. H., and Guo, H. C. (2024). Does AI-driven technostress promote or hinder employees' artificial intelligence adoption intention? A moderated mediation model of affective reactions and technical self-efficacy. *Psychol. Res. Behav. Manag.* 17, 413–427. doi: 10.2147/prbm.S441444

Cheung, G. W., Cooper-Thomas, H. D., Lau, R. S., and Wang, L. C. (2024). Reporting reliability, convergent and discriminant validity with structural equation modeling: a review and best-practice recommendations. *Asia Pac. J. Manag.* 41, 745–783. doi: 10.1007/s10490-023-09871-y

Chiu, T. K. F., and Chai, C. S. (2020). Sustainable curriculum planning for artificial intelligence education: a self-determination theory perspective. *Sustainability* 12:5568. doi: 10.3390/su12145568

Chiu, T. K. F., Falloon, G., Song, Y. J., Wong, V. W. L., Zhao, L., and Ismailov, M. (2024). A self-determination theory approach to teacher digital competence development. *Comput. Educ.* 214:Article 105017. doi: 10.1016/j.compedu.2024.105017

Cho, H., and Jung, H. (2025). Job autonomy and job satisfaction of crowdworkers: mediating effects of self-efficacy and meaningfulness of work. *SAGE Open* 15:21582440251328488. doi: 10.1177/21582440251328488

Dahri, N. A., Yahaya, N., Al-Rahmi, W. M., Aldraiweesh, A., Alturki, U., Almutairy, S., et al. (2024). Extended TAM based acceptance of AI-powered ChatGPT for supporting metacognitive self-regulated learning in education: a mixed-methods study. *Heliyon* 10:e29317. doi: 10.1016/j.heliyon.2024.e29317

Dash, G., and Paul, J. (2021). CB-SEM vs. PLS-SEM methods for research in social sciences and technology forecasting. *Technol. Forecast. Soc. Change* 173:121092. doi: 10.1016/j.techfore.2021.121092

Deci, E. L., Olafsen, A. H., and Ryan, R. M. (2017). Self-determination theory in work organizations: the state of a science. *Annu. Rev. Organ. Psych. Organ. Behav.* 4, 19–43. doi: 10.1146/annurev-orgpsych-032516-113108

Deci, E. L., and Ryan, R. M. (2000). The "what" and "why" of goal pursuits: human needs and the self-determination of behavior. *Psychol. Inq.* 11, 227–268. doi: 10.1207/s15327965pli1104_01

Dilek, M., Baran, E., and Aleman, E. (2025). AI literacy in teacher education: empowering educators through critical co-discovery. *J. Teach. Educ.* 76, 294–311. doi: 10.1177/00224871251325083

Ding, A. C. E., Shi, L. H., Yang, H. T., and Choi, I. (2024). Enhancing teacher AI literacy and integration through different types of cases in teacher professional development. *Comput. Educ. Open* 6:100178. doi: 10.1016/j.caeo.2024.100178

Dorenkamp, I., and Ruhle, S. (2019). Work-life conflict, professional commitment, and job satisfaction among academics. *J. High. Educ.* 90, 56–84. doi: 10.1080/00221546.2018.1484644

Ertiö, T., Eriksson, T., Rowan, W., and McCarthy, S. (2024). The role of digital leaders' emotional intelligence in mitigating employee technostress. *Bus. Horiz.* 67, 399–409. doi: 10.1016/j.bushor.2024.03.004

Fotiadis, A., Abdulrahman, K., and Spyridou, A. (2019). The mediating roles of psychological autonomy, competence and relatedness on work-life balance and well-being. *Front. Psychol.* 10:1267. doi: 10.3389/fpsyg.2019.01267

Francis, S. P., Kolil, V. K., Pavithran, V., Ray, I., and Achuthan, K. (2024). Exploring gender dynamics in cybersecurity education: a self-determination theory and social cognitive theory perspective. *Comput. Secur.* 144:103968. doi: 10.1016/j.cose.2024.103968

Gagné, M., Parent-Rocheleau, X., Bujold, A., Gaudet, M. C., and Lirio, P. (2022). How algorithmic management influences worker motivation: a self-determination theory perspective. *Can. Psychol.* 63, 247–260. doi: 10.1037/cap0000324

Greenhaus, J. H., and Beutell, N. J. (1985). Sources of conflict between work and family roles. *Acad. Manag. Rev.* 10, 76–88. doi: 10.2307/258214

Hashem, R., Ali, N., El Zein, F., Fidalgo, P., and Abu Khurma, O. (2024). Ai to the rescue: exploring the potential of ChatGPT as a teacher ally for workload relief and burnout prevention. *Res. Pract. Technol. Enhanc. Learn.* 19:23. doi: 10.58459/rptel.2024.19023

Haw, J. Y., Nalipay, M. J. N., and King, R. B. (2024). Perceived relatedness-support matters most for teacher well-being: a self-determination theory perspective. *Teach. Teach.* 30, 745–761. doi: 10.1080/13540602.2023.2263736

Hofer, S. I., Nistor, N., and Scheibenzuber, C. (2021). Online teaching and learning in higher education: lessons learned in crisis situations. *Comput. Hum. Behav.* 121:106789. doi: 10.1016/j.chb.2021.106789

Hwang, G.-J., Xie, H., Wah, B. W., and Gašević, D. (2020). Vision, challenges, roles and research issues of artificial intelligence in education. *Comput. Educ.* 1:100001. doi: 10.1016/j.caeai.2020.100001

Inigo, M., and Raufaste, E. (2019). Academics' motivations explain time-allocation and well-being at work. *Eur. Rev. Appl. Psychol.* 69, 19–30. doi: 10.1016/j.erap.2018.11.002

Janssen, S., van Vuuren, M., and de Jong, M. D. T. (2013). Identifying support functions in developmental relationships: a self-determination perspective. *J. Vocat. Behav.* 82, 20–29. doi: 10.1016/j.jvb.2012.09.005

Ji, Y., Zhong, M. X., Lyu, S., Li, T. T., Niu, S. J., and Zhan, Z. H. (2025). How does AI literacy affect individual innovative behavior: the mediating role of psychological need satisfaction, creative self-efficacy, and self-regulated learning. *Educ. Inf. Technol.* 30, 16133–16162. doi: 10.1007/s10639-025-13437-4

Jose, B., Cherian, J., Verghis, A. M., Varghise, S. M., Mumthas, S., and Joseph, S. (2025). The cognitive paradox of AI in education: between enhancement and erosion. *Front. Psychol.* 16:1550621. doi: 10.3389/fpsyg.2025.1550621

Kaliisa, R., Rienties, B., Morch, A., and Kluge, A. (2022). Social learning analytics in computer-supported collaborative learning environments: a systematic review of empirical studies. *Comput. Educ. Open* 3:100073. doi: 10.1016/j.caeo.2022.100073

Kelley, M., and Wenzel, T. (2025). Advancing artificial intelligence literacy in teacher education through professional partnership inquiry. *Educ. Sci.* 15:659. doi: 10.3390/educsci15060659

Khawand, S., and Zargar, P. (2022). Job autonomy and work-life conflict: a conceptual analysis of teachers' wellbeing during COVID-19 pandemic. *Front. Psychol.* 13:2848. doi: 10.3389/fpsyg.2022.882848

Kim, J. (2024). Leading teachers' perspective on teacher-AI collaboration in education. *Educ. Inf. Technol.* 29, 8693–8724. doi: 10.1007/s10639-023-12109-5

Klassen, R. M., and Chiu, M. M. (2010). Effects on teachers' self-efficacy and job satisfaction: teacher gender, years of experience, and job stress. *J. Educ. Psychol.* 102, 741–756. doi: 10.1037/a0019237

Klassen, R. M., Perry, N. E., and Frenzel, A. C. (2012). Teachers' relatedness with students: an underemphasized component of teachers' basic psychological needs. *J. Educ. Psychol.* 104, 150–165. doi: 10.1037/a0026253

Lai, C., and Jin, T. (2021). Teacher professional identity and the nature of technology integration. *Comput. Educ.* 175:104314. doi: 10.1016/j.compedu.2021.104314

Landolfi, A., Barattucci, M., De Rosa, A., and Lo Presti, A. (2021). The Association of job and Family Resources and demands with life satisfaction through work-family balance: a longitudinal study among Italian schoolteachers during the COVID-19 pandemic. *Behav. Sci.* 11:136. doi: 10.3390/bs11100136

Laupichler, M. C., Aster, A., Schirch, J., and Raupach, T. (2022). Artificial intelligence literacy in higher and adult education: a scoping literature review. *Comput. Educ.* 3:100101. doi: 10.1016/j.caeai.2022.100101

Layek, D., and Koodamara, N. K. (2024). Motivation, work experience, and teacher performance: a comparative study. *Acta Psychol.* 245:104217. doi: 10.1016/j.actpsy.2024.104217

Li, L. N., Kanchanapoom, K., Deeprasert, J., Duan, N. G., and Qi, Z. (2025). Unveiling the factors shaping teacher job performance: exploring the interplay of personality traits, perceived organizational support, self-efficacy, and job satisfaction. *BMC Psychol.* 13:14. doi: 10.1186/s40359-024-02324-1

Li, M., and Yu, Z. G. (2022). Teachers' satisfaction, role, and digital literacy during the COVID-19 pandemic. *Sustainability* 14:1121. doi: 10.3390/su14031121

Lin, T. Q., Zhang, J. Y., and Xiong, B. (2025). Effects of technology perceptions, teacher beliefs, and AI literacy on AI technology adoption in sustainable mathematics education. *Sustainability* 17:3698. doi: 10.3390/su17083698

Lyu, Y. J., and Zhu, H. (2019). The predictive effects of workplace ostracism on employee attitudes: a job embeddedness perspective. *J. Bus. Ethics* 158, 1083–1095. doi: 10.1007/s10551-017-3741-x

Ma, T. N., and Vu, H. V. (2024). How job autonomy influences life satisfaction: the role of work-family enrichment and segmentation preference. *Int. J. Organ. Anal.* 32, 1833–1849. doi: 10.1108/ijoa-05-2023-3762

Meharunisa, S., Almugren, H., Sarabdeen, M., Mabrouk, F., and Kijas, A. C. M. (2024). The impact of artificial intelligence on women's empowerment, and work-life balance in Saudi educational institutions. *Front. Psychol.* 15:2541. doi: 10.3389/fpsyg.2024.1432541

Molefi, R. R., Ayanwale, M. A., Kurata, L., and Chere-Masopha, J. (2024). Do in-service teachers accept artificial intelligence-driven technology? The mediating role of school support and resources. *Comp. Educ. Open* 6:100191. doi: 10.1016/j.caeo.2024.100191

Mulyani, S., Salameh, A. A., Komariah, A., Timoshin, A., Hashim, N., Fauziah, R. S. P., et al. (2021). Emotional regulation as a remedy for teacher burnout in special schools: evaluating school climate, teacher's work-life balance and children behavior. *Front. Psychol.* 12:655850. doi: 10.3389/fpsyg.2021.655850

Naidoo, L., and Wagner, S. (2020). Thriving, not just surviving: the impact of teacher mentors on pre-service teachers in disadvantaged school contexts. *Teach. Teach. Educ.* 96:103185. doi: 10.1016/j.tate.2020.103185

Ng, D. T. K., Leung, J. K. L., Su, J. H., Ng, R. C. W., and Chu, S. K. W. (2023). Teachers' AI digital competencies and twenty-first century skills in the post-pandemic world. *Educ. Technol. Res. Dev.* 71, 137–161. doi: 10.1007/s11423-023-10203-6

Ning, Y. M., Zhang, W. J., Yao, D. M., Fang, B. W., Xu, B. Y., and Wijaya, T. T. (2025). Development and validation of the artificial intelligence literacy scale for teachers (AILST). *Educ. Inf. Technol.* 30, 17769–17803. doi: 10.1007/s10639-025-13347-5

Ozudogru, G., and Durak, H. Y. (2025). Conceptualizing pre-service teachers' artificial intelligence readiness and examining its relationship with various variables: the role of artificial intelligence literacy, digital citizenship, artificial intelligence-enhanced innovation and perceived threats from artificial intelligence. *Inf. Dev.* doi: 10.1177/02666669251335657

Pace, F., D'Urso, G., Zappulla, C., and Pace, U. (2021). The relation between workload and personal well-being among university professors. *Curr. Psychol.* 40, 3417–3424. doi: 10.1007/s12144-019-00294-x

Pan, X. (2020). Technology acceptance, technological self-efficacy, and attitude toward technology-based self-directed learning: learning motivation as a mediator. *Front. Psychol.* 11:564294. doi: 10.3389/fpsyg.2020.564294

Pasamar, S., Johnston, K., and Tanwar, J. (2020). Anticipation of work-life conflict in higher education. *Employ. Relat.* 42, 777–797. doi: 10.1108/er-06-2019-0237

Pauw, L. S., Sauter, D. A., van Kleef, G. A., Lucas, G. M., Gratch, J., and Fischer, A. H. (2022). The avatar will see you now: support from a virtual human provides socio-emotional benefits. *Comput. Hum. Behav.* 136:107368. doi: 10.1016/j.chb.2022.107368

Pelau, C., Dabija, D. C., and Ene, I. (2021). What makes an AI device human-like? The role of interaction quality, empathy and perceived psychological anthropomorphic characteristics in the acceptance of artificial intelligence in the service industry. *Comput. Human Behav.* 122:Article 106855. doi: 10.1016/j.chb.2021.106855

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, 879–903. doi: 10.1037/0021-9010.88.5.879

Prasad, C. V., and Pasupathi, G. (2025). Exploring the influence of work-life balance and work environment on job satisfaction among college teachers. *Public Organ. Rev.* doi: 10.1007/s11115-024-00807-8

Prenger, R., Poortman, C. L., and Handelzalts, A. (2017). Factors influencing teachers' professional development in networked professional learning communities. *Teach. Teach. Educ.* 68, 77–90. doi: 10.1016/j.tate.2017.08.014

Racero, F. J., Bueno, S., and Gallego, M. D. (2020). Predicting students' behavioral intention to use open source software: a combined view of the technology acceptance model and self-determination theory. *Appl. Sci. Basel* 10:Article 2711. doi: 10.3390/app10082711

Rahimi, H., Hejazi, S. Y., Lou, N. M., and Heidarzadeh, M. (2024). Are teachers with better quality of work life more innovative? The mediating roles of psychological empowerment and teaching mindsets. *Acta Psychol.* 247:104315. doi: 10.1016/j.actpsy.2024.104315

Raykov, T., and Marcoulides, G. A. (2019). Thanks coefficient alpha, we still need you! *Educ. Psychol. Meas.* 79, 200–210. doi: 10.1177/0013164417725127

Ryan, R. M., and Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *Am. Psychol.* 55, 68–78. doi: 10.1037/0003-066x.55.1.68

Ryan, R. M., and Deci, E. L. (2020). Intrinsic and extrinsic motivation from a self-determination theory perspective: definitions, theory, practices, and future directions. *Contemp. Educ. Psychol.* 61:101860. doi: 10.1016/j.cedpsych.2020.101860

Scherer, R., Siddiq, F., and Tondeur, J. (2019). The technology acceptance model (TAM): a meta-analytic structural equation modeling approach to explaining teachers' adoption of digital technology in education. *Comput. Educ.* 128, 13–35. doi: 10.1016/j.compedu.2018.09.009

Shen, Y. Y., and Cui, W. C. (2024). Perceived support and AI literacy: the mediating role of psychological needs satisfaction. *Front. Psychol.* 15:1415248. doi: 10.3389/fpsyg.2024.1415248

Şimşek, T. (2025). "Human-machine collaboration: how do employees interact with artificial intelligence?" in AI use in social sciences. (Hershey, PA: IGI Global Scientific Publishing), 157–180.

Singh, H. P., and Aziz, A. A. (2025). Impact of intelligent learning assistants on creativity of university students: a self-determination theory perspective. *Future Bus. J.* 11:122. doi: 10.1186/s43093-025-00544-4

Singh, P., Bala, H., Dey, B. L., and Filieri, R. (2022). Enforced remote working: the impact of digital platform-induced stress and remote working experience on technology exhaustion and subjective wellbeing. *J. Bus. Res.* 151, 269–286. doi: 10.1016/j.jbusres.2022.07.002

Sperling, K., Stenberg, C. J., McGrath, C., Akerfeldt, A., Heintz, F., and Stenliden, L. (2024). In search of artificial intelligence (AI) literacy in teacher education: a scoping review. *Comput. Educ. Open* 6:100169. doi: 10.1016/j.caeo.2024.100169

Sumak, B., Hericko, M., and Pusnik, M. (2011). A meta-analysis of e-learning technology acceptance: the role of user types and e-learning technology types. *Comput. Hum. Behav.* 27, 2067–2077. doi: 10.1016/j.chb.2011.08.005

Teo, T. (2011). Factors influencing teachers' intention to use technology: model development and test. *Comput. Educ.* 57, 2432–2440. doi: 10.1016/j.compedu.2011.06.008

Teo, T., and Noyes, J. (2014). Explaining the intention to use technology among pre-service teachers: a multi-group analysis of the unified theory of acceptance and use of technology. *Interact. Learn. Environ.* 22, 51–66. doi: 10.1080/10494820.2011.641674

Troesch, L. M., and Bauer, C. E. (2017). Second career teachers: job satisfaction, job stress, and the role of self-efficacy. *Teach. Teach. Educ.* 67, 389–398. doi: 10.1016/j.tate.2017.07.006

Turk, M., Heddy, B. C., and Danielson, R. W. (2022). Teaching and social presences supporting basic needs satisfaction in online learning environments: how can presences and basic needs happily meet online? *Comput. Educ.* 180:Article 104432. doi: 10.1016/j.compedu.2022.104432

Van den Broeck, A., Ferris, D. L., Chang, C. H., and Rosen, C. C. (2016). A review of self-determination theory's basic psychological needs at work. *J. Manage.* 42, 1195–1229. doi: 10.1177/0149206316632058

Venkatesh, V., and Davis, F. D. (2000). A theoretical extension of the technology acceptance model: four longitudinal field studies. *Manag. Sci.* 46, 186–204. doi: 10.1287/mnsc.46.2.186.11926

Venkatesh, V., Morris, M. G., Davis, G. B., and Davis, F. D. (2003). User acceptance of information technology: toward a unified view. *MIS Q.* 27, 425–478. doi: 10.2307/30036540

Wang, K., Cui, W. C., and Yuan, X. (2025). Artificial intelligence in higher education: the impact of need satisfaction on artificial intelligence literacy mediated by self-regulated learning strategies. *Behav. Sci.* 15:Article 165. doi: 10.3390/bs15020165

Wang, H., Lee, S. Y., and Hall, N. C. (2022). Coping profiles among teachers: implications for emotions, job satisfaction, burnout, and quitting intentions. *Contemp. Educ. Psychol.* 68:102030. doi: 10.1016/j.cedpsych.2021.102030

Wang, B. C., Rau, P. L. P., and Yuan, T. Y. (2023). Measuring user competence in using artificial intelligence: validity and reliability of artificial intelligence literacy scale. *Behav. Inf. Technol.* 42, 1324–1337. doi: 10.1080/0144929x.2022.2072768

Wei, C. W., and Ye, J. H. (2022). The impacts of work-life balance on the emotional exhaustion and well-being of college teachers in China. *Healthcare* 10:2234. doi: 10.3390/healthcare10112234

Wu, T. J., Pan, Q. H., Zhang, K. X., and Zhang, R. X. (2024). Flexible working on employees' quality of work life: the role of locus of control through psychological perspective. *Curr. Psychol.* 43, 33307–33322. doi: 10.1007/s12144-024-06848-y

Xia, Q., Chiu, T. K. F., Chai, C. S., and Xie, K. (2023). The mediating effects of needs satisfaction on the relationships between prior knowledge and self-regulated learning through artificial intelligence chatbot. *Br. J. Educ. Technol.* 54, 967–986. doi: 10.1111/bjet.13305

Xia, Q., Chiu, T. K. F., Lee, M., Sanusi, I. T., Dai, Y., and Chai, C. S. (2022). A self-determination theory (SDT) design approach for inclusive and diverse artificial intelligence (AI) education. *Comput. Educ.* 189:104582. doi: 10.1016/j.compedu.2022.104582

Xiao, J., Yang, Y. L., and Li, M. (2025). Empirical study on the feasibility of hybrid-flexible training model for developing teachers' artificial intelligence competence. *Educ. Inf. Technol.* 30, 16835–16860. doi: 10.1007/s10639-025-13460-5

Yang, Y., and Ling, Q. (2023). The influence of existence-relatedness-growth need satisfaction and job burnout of young university teachers: the mediating role of job satisfaction. *Front. Psychol.* 14:5742. doi: 10.3389/fpsyg.2023.1205742

Yildirim, M., Dilekçi, U., and Manap, A. (2024). Mediating roles of meaning in life and psychological flexibility in the relationships between occupational stress and job satisfaction, job performance, and psychological distress in teachers. *Front. Psychol.* 15:1349726. doi: 10.3389/fpsyg.2024.1349726

Yin, H. B., and Huang, S. H. (2021). Applying structural equation modelling to research on teaching and teacher education: looking back and forward. *Teach. Teach. Educ.* 107:103438. doi: 10.1016/j.tate.2021.103438

Yu, H. (2024). The application and challenges of ChatGPT in educational transformation: new demands for teachers' roles. *Heliyon* 10:4289. doi: 10.1016/j.heliyon.2024.e24289

Zagni, B., Pellegrino, G., Ianes, D., and Scrimin, S. (2025). Unraveling teacher stress: a cumulative model of risks and protective factors in Italian schools. *Int. J. Educ. Res.* 131:Article 102603. doi: 10.1016/j.ijer.2025.102603

Zawacki-Richter, O., Marin, V. I., Bond, M., and Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education – where are the educators? *Int. J. Educ. Technol. High. Educ.* 16:39. doi: 10.1186/s41239-019-0171-0

Zhang, Z. (2023). The impact of the artificial intelligence industry on the number and structure of employments in the digital economy environment. *Technol. Forecast. Soc. Chang.* 197:122881. doi: 10.1016/j.techfore.2023.122881

Zheng, J. Q., Cheung, K. C., Sit, P. S., and Lam, C. C. (2025). Unfolding key factors of resilience in ICT cognitive-motivational engagement: global evidence from machine learning techniques. *Int. J. Educ. Res.* 131:102607. doi: 10.1016/j.ijer.2025.102607

Zheng, J. S., and Zhang, T. (2025). Association between AI awareness and emotional exhaustion: the serial mediation of job insecurity and work interference with family. *Behav. Sci.* 15:401. doi: 10.3390/bs15040401

Zhou, J. S., Shen, L. H., and Chen, W. B. (2024). How ChatGPT transformed teachers: the role of basic psychological needs in enhancing digital competence. *Front. Psychol.* 15:1458551. doi: 10.3389/fpsyg.2024.1458551

Zhou, T., Wu, X. L., Wang, Y. D., Wang, Y. L., and Zhang, S. A. (2024). Application of artificial intelligence in physical education: a systematic review. *Educ. Inf. Technol.* 29, 8203–8220. doi: 10.1007/s10639-023-12128-2

frontiers | Frontiers in Psychology

TYPE Opinion
PUBLISHED 07 October 2025
DOI 10.3389/fpsyg.2025.1659272

OPEN ACCESS

EDITED BY
Tour Liu,
Tianjin Normal University, China

REVIEWED BY
Magda Julissa Rojas-Bahamón,
Jorge Eliécer Gaitán, Colombia
Grace Jochan,
St. Joseph's University, India

*CORRESPONDENCE
Sobi Thomas
✉ sobykannalil@gmail.com

RECEIVED 03 July 2025
ACCEPTED 19 September 2025
PUBLISHED 07 October 2025

CITATION
Thomas S and Manalil P (2025) Digital silence: the psychological impact of being shadow banned on mental health and self-perception. *Front. Psychol.* 16:1659272.
doi: 10.3389/fpsyg.2025.1659272

Digital silence: the psychological impact of being shadow banned on mental health and self-perception

Sobi Thomas[1,2]* and Paul Manalil[1]

[1]School of Social Science, Arts and Humanities, Lincoln University College, Petaling Jaya, Malaysia,
[2]Department of Communication and Media Studies, Marian College Kuttikkanam, Kuttikkanam, India

KEYWORDS

shadow banning, algorithmic invisibility, self-perception, digital silence, mental health, media psychology, online validation, emotional regulation

1 Introduction: when the feed goes quiet

Imagine that you publish a well-thought-out post, photo, or video—only to watch it silently sink into obscurity. No likes, no comments, no shares. At first, you can brush it off as a fluke or a bad content day. But for many of us, specifically for those who are on active social media platforms like Instagram, Tiktok, or X (previously Twitter), Shadow banning could be the primarily cause resulting in their silence. Shadow banning could be defined as an algorithmic hiding in which the content is quietly de-amplified without no indication (Liu et al., 2023). To differentiate overt censorship from shadow banning, it is an act with a conscious face whereas shadow banning is invisible and creates a sense of social erasure that could potential result in emotional disorientation and psychological distress. Recently, research studies have begun emphasizing the importance to recognize the shadow banning not only as a technical limitation but also on a broader spectrum on digital exclusion and algorithmic marginalization (Delmonaco et al., 2024).

In this paper, we examine shadow banning more as an intensely subjective psycho-existential phenomenon rather than as a technical bug or policy enforcement strategy. Findings of this study show that Shadow banning emotionally affect the self-concept leading to disruptions in digital social feedbacks. The individuals are therefore compelled to rely for validation identification, reinforcement, and social inclusion. This study did a detailed analysis of the literature in media psychology and theories of emotional and digital behavior, and concludes that non-transparency of the social media platforms causes distress of individuals, and it needs to be addressed urgently.

2 Understanding shadow banning and its affective mechanism

Shadow banning also known as Stealth banning, silently prevents or restricts a user's reach in the social media platforms. It is a kind of algorithmic suppression without suspending the account. Unaware of the invisibility of the post in the community the user till continues posting, but the message never appears in search results, hashtags, or regular feeds, leading to decreased engagement. These users are, in fact, speaking to a void. This digital silence can be described as a vocal message within the social media economy.

Feedback is a sustainable rejuvenating factor of the online platforms. The activating responses through "likes, comments, reposts and follows" are emotional assets which indicates self-affirmation. These validations cannot be ignored as they activate neural centres which releases dopamine. When these signals vanish into thin air with no indication of why the users feel lost, rejected and struggle with cognitive dissonance (Politte-Corn et al., 2024). Am I ignored? Is my content awful? Have I done something inappropriate? The withdrawal from engagement is a psychological riddle that upsets the self-worth. From a psychological standpoint, this dynamic activates the mesolimbic dopamine system, reinforcing the role of social affirmation in self-perception (Cross et al., 2025). Cognitive dissonance arises when one's self-image as a socially engaged digital citizen clashes with unexplained algorithmic suppression. A qualitative analysis using Impression Management Theory and Cognitive Dissonance Theory found that teens experience dissonance when their social media presence conflicts with their real-world identity, often leading to discomfort and eventual withdrawal from online activity (Marta and Miletresia, 2022).

The line graph in the Figure 1 illustrates a noticeable decline in user engagement (likes, comments, shares) following a suspected shadow ban. The data is based on user-reported case studies, showing normal interaction patterns in the days prior (Days 1–15), followed by a significant drop post-event (Days 16–30). This pattern exemplifies the experience of "digital silence," where content visibility is algorithmically suppressed without user notification, leading to emotional confusion and self-doubt. While this visual is based on informal reports and lacks formal statistical validation, it reflects a recurring pattern documented across multiple user narratives.

3 Emotional dysregulation and self-doubt in a platformed identity

Online, the identity is not just described—it is staged and legitimated in the public sphere. The self is algorithmically discernible, constituted with interaction metrics and validation from followers. When a user is shadow banned, they are systematically excluded from the social world. The shock invisibility disrupts emotional regulatory protocols and can induce depression symptoms, anxiety, and compulsive checking of content behaviors (Wikman et al., 2022).

The concerns of social exclusion were studied by media psychologists in recent years and their findings focus on the "indefiniteness" of shadow banning. The users were not told about the banning and the indefinite nature of such banning. The individuals quite often doubt their perception of reality and the emotional cost of exclusion from the social media platforms in high, particularly for creators of activist postings, often associated with political assertions of minority users (Powers et al., 2013). The freedom of expression of such communities is infringed through shadow banning. As no one is held responsible it makes emotional recuperation more difficult. The lack of feedback from the social media platforms, particularly among the users result in emotional dysregulation or a difficulty in managing emotional responses in accordance with the contextual demands (Rogier et al., 2024).

For the individual users the silent platforms are a failure of their own. Such instances ultimately lead to detrimental thinking patterns like repeated checking of the reach of the posts, resubmission and republishing of posts or immerse in self-critical thinking. It not only frustrates but psychologically damage the user (Da Silva Pinho et al., 2024).

4 Algorithmic inequality and emotional toll of shadow banning

The impact of shadow banning is not equally affected. The posts which are themed on sexuality, racial disturbances, social activism or body-positive are invariably censored. When these posts are not against the rules it reaches the users (Foster et al., 2021). There are many inherent structural inequalities due to algorithmic governance.

The subaltern and fringe groups in the society who are considered marginalised population always feel that their visibility is conditional and carefully crafted. The content provides belonging to queer and fat rights organisers negotiate their own space in the media for interactions and survival protests (Escobar-Viera et al., 2023). Some minority groups like queer had modest following on Instagram, but later when they discussed other general social issues there was sharp drop in views on all subsequent posts. The digital silencing occurs without formal notices and eventually it leads to distress and a temporary social media hiatus; an emotional erasure that sustains systemic silencing. It is a shame on individuals who feel that invisibility is a personal failure than a structure defect of media. As Covin (2021) emphasizes, shadow banning can lead to "unseen shame," where users privately struggle with feelings of inadequacy, internalizing their online invisibility as a personal failing, despite the lack of explicit criticism from others.

Recent studies on digital exclusion reveal that algorithmic decisions can perpetuate existing social inequalities online, leaving users feeling unfairly penalized for their identity or views. The constant pressure to create content, coupled with the algorithm's silent devaluation of their voice, can be exhausting (Nair et al., 2024).

5 Shadow banning stems from inherent ambiguity?

When uncertainty increases anxiety and causes psychological distress it eventually leads to repetitive negative thoughts and thereby aggravate mental health concerns (Altan-Atalay et al., 2023).

The shadow banned users repeatedly fall into uncertainties even as they continue the futile exercise of selecting hashtags. The emotional exhaustion produces helplessness and bewilderment. The ambiguity linked to the posting in the social media can impact on trans-diagnostic factors linked to anxiety disorders and obsessive rumination. It renders the users more susceptible to distress (Pinciotti et al., 2021). The intolerant situation caused by uncertainty compels the users to quit the site because silence became unsustainable psychologically. Covin (2021) notes that this hidden shame in digital environments rarely has a reintegrated

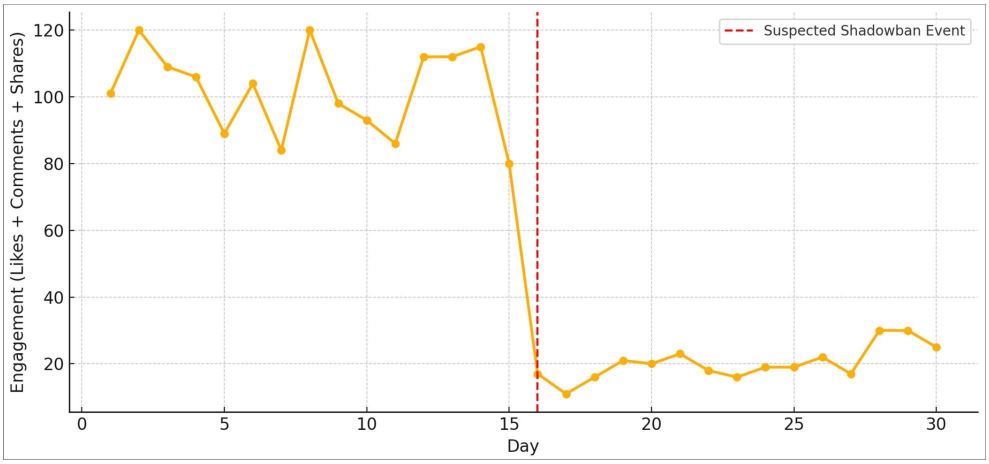


FIGURE 1
Sudden drop in engagement metrics after suspected shadowban event. This figure is based on a composite of self-reported case patterns drawn from user forums and anecdotal experiences. It is presented illustratively to depict a typical engagement trajectory following suspected shadow banning.

function. It isolates the user and increases his or her isolation. This corresponds with Jochan and Banerjee (2021) argument that shame in digital environments rarely has a reintegrated function; instead, it isolates the individual and deepens alienation.

The obscure element in the shadow banning process disrupts digital trust. Though the social media platforms claim freedom of expression they involve in stealth moderation that facilitates self-censorship and self-policing (Wang and Kim, 2023). This phenomenon can subtly persuade unwilling users into altering their tone and the themes, which eventually lead to emotional conformity due to prolonged limitation on the freedom of expression. The present study focuses on the urgent need for specific interventions to address the issues of ambiguity and emotional impact of algorithmic governance related to shadow banning. The negative psychological effects are far-reaching and it includes exclusion, shame and loss of trust. The transparency in the process of algorithmic governance and alleviation of deeply emotional and identity related constraints the users face online must be prioritized in finding solutions (Risius and Blasiak, 2024).

6 A humane platform design and emotional transparency needed

There is an invisible layer of shame in the social media platforms which highlights not only the fundamental issues of algorithmic transparency, but also the hidden psychological costs, ensuring that design responses attend to both external visibility and internal will-being (Covin, 2021). The social media platforms must acknowledge the damage caused by opaque algorithms and adopt transparent practices to reduce the emotional harm done to the users. If the reasons behind the content moderation decisions are explained the platforms can reduce user anxiety and build trust, creating a more open and reliable online environment (Jansen and Krämer, 2023).

Platforms should design with users' mental health in mind, incorporating features such as notifications, appeal options, and transparent explanations for content visibility. Fair governance demands transparency, due process, and accountability, rather than unexplained penalties (Russ et al., 2014). Openness is not a technical remedy; it is a psycho logical necessity.

The mental health practitioners should include algorithmic exclusion within their conceptual framework of digital trauma (Barton et al., 2023). The sudden invisibility resulting from shadow banning can precipitate profound identity crises and emotional distress. Mental health professionals should be trained to address these concerns. Moreover, media literacy initiatives should extend beyond filter bubbles and misinformation to encompass the emotional consequences of algorithmic silence. Further research is warranted to explore the intersections between online trauma and other digital harms, such as cyberbullying, harassment, and community disintegration, to comprehensively understand the phenomenon's scope and implications (Delmonaco et al., 2024).

Despite being dismissed as conspiracy theories, shadow banning can cause real harm. We need more research that combines platform data, user experiences, and signs of psychological distress to understand the true mental health impact of being algorithmically suppressed online.

7 Conclusion: making the invisible visible

Visibility is validation in the social media platforms. Shadow banning turns invisibility into a weapon, and the silent treatment of the feed a tool of emotional coercion. Faith in the platforms erodes, shattering the users' perceptions of the self, and digital neurosis and self-doubt intensify (Van Noordt et al., 2015).

This opinion piece contends that shadow banning transcends content moderation, posing a significant psychological concern. By disrupting emotional regulation, exacerbating social inequalities, and fostering cognitive dissonance, it takes a profound toll on users. To mitigate this, media platforms must prioritize the emotional impact of algorithmic governance, lest users continue to experience silent suffering, overshadowed by both code and emotional distress.

To make the invisible visible is the first step toward justice—technical, social, and psychological. Let that apply not only to content, but to the human costs hidden behind the feed.

Author contributions

ST: Conceptualization, Validation, Visualization, Writing – original draft, Writing – review & editing. PM: Supervision, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that no financial support was received for the research and/or publication of this article.

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.

Generative AI statement

The author(s) declare that no Gen AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

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

Altan-Atalay, A., Tuncer, I., King, N., Önol, B., Sözeri, Y., and Tezel, S. (2023). Repetitive negative thinking during ambiguous situations: interactive roles of looming cognitive style and intolerance of uncertainty. *J. Behav. Ther. Exp. Psychiatry* 79:101840. doi: 10.1016/j.jbtep.2023.101840

Barton, B. B., Reinhard, M. A., Goerigk, S., Wüstenberg, T., Musil, R., Ehring, T., et al. (2023). Association between the behavioral response during social exclusion and recalled childhood maltreatment. *Behav. Res. Ther.* 160:104232. doi: 10.1016/j.brat.2022.104232

Covin, R. (2021). *The Private Shame of Social Media Users*. Medium. Available online at: https://rogercovin.medium.com/internet-shame-a-common-and-unseen-private-experience-5bccebf757a (Accessed September 15, 2025).

Cross, E. A., Borland, J. M., Shaughnessy, E. K., Lee, S. D., Vu, V., Sambor, E. A., et al. (2025). Distinct subcircuits within the mesolimbic dopamine system encode the salience and valence of social stimuli. *Psychopharmacology* 242, 2219–2232. doi: 10.1007/s00213-025-06793-z

Da Silva Pinho, A., Céspedes Izquierdo, V., Lindström, B., and Van Den Bos, W. (2024). Youths' sensitivity to social media feedback: a computational account. *Sci. Adv.* 10:eadp8775. doi: 10.1126/sciadv.adp8775

Delmonaco, D., Mayworm, S., Thach, H., Guberman, J., Augusta, A., and Haimson, O. L. (2024). "What are you doing, TikTok?": how marginalized social media users perceive, theorize, and "prove" shadowbanning. *Proc. ACM Hum. Comput. Interact.* 8, 1–39. doi: 10.1145/3637431

Escobar-Viera, C. G., Porta, G., Coulter, R. W. S., Martina, J., Goldbach, J., and Rollman, B. L. (2023). A chatbot-delivered intervention for optimizing social media use and reducing perceived isolation among rural-living LGBTQ+ youth: development, acceptability, usability, satisfaction, and utility. *Internet Interv.* 34:100668. doi: 10.1016/j.invent.2023.100668

Foster, M. D., Tassone, A., and Matheson, K. (2021). Tweeting about sexism motivates further activism: a social identity perspective. *Br. J. Soc. Psychol.* 60, 741–764. doi: 10.1111/bjso.12431

Jansen, M.-P., and Krämer, N. C. (2023). Balancing perceptions of targeting: an investigation of political microtargeting transparency through a calculus approach. *PLoS ONE* 18:e0295329. doi: 10.1371/journal.pone.0295329

Jochan, G. M., and Banerjee, T. (2021). "Shaming in the internet era: evaluating the reintegrative function of shame in digital spaces," in *Shame 4.0*, eds. C.-H. Mayer, E. Vanderheiden, and P. T. P. Wong (Cham: Springer International Publishing), 431–454.

Liu, J., Lee, D. N., and Stevens, E. M. (2023). Characteristics associated with young adults' intentions to engage with anti-vaping Instagram posts. *Int. J. Environ. Res. Public Health* 20:6054. doi: 10.3390/ijerph20116054

Marta, R. F., Miletresia, and Fernandes, M. (2022). Reflections of today's teens behavior: from impressions on social media to cognitive dissonance. *Salus Cultura* 2, 81–91. doi: 10.55480/saluscultura.v2i1.49

Nair, K., Mosleh, M., and Kouchaki, M. (2024). Racial minorities face discrimination from across the political spectrum when seeking to form ties on social media: evidence from a field experiment. *Psychol. Sci.* 35, 1278–1286. doi: 10.1177/09567976241274738

Pinciotti, C. M., Riemann, B. C., and Abramowitz, J. S. (2021). Intolerance of uncertainty and obsessive-compulsive disorder dimensions. *J. Anxiety Disord.* 81:102417. doi: 10.1016/j.janxdis.2021.102417

Politte-Corn, M., Pegg, S., Dickey, L., and Kujawa, A. (2024). Neural reactivity to social reward moderates the association between social media use and momentary positive affect in adolescents. *Affect. Sci.* 5, 281–294. doi: 10.1007/s42761-024-00237-1

Powers, K. E., Somerville, L. H., Kelley, W. M., and Heatherton, T. F. (2013). Rejection sensitivity polarizes striatal–medial prefrontal activity when anticipating social feedback. *J. Cogn. Neurosci.* 25, 1887–1895. doi: 10.1162/jocn_a_00446

Risius, M., and Blasiak, K. M. (2024). Shadowbanning: an opaque form of content moderation. *Bus. Inf. Syst. Eng.* 66, 817–829. doi: 10.1007/s12599-024-00905-3

Rogier, G., Muzi, S., and Pace, C. S. (2024). Social media misuse explained by emotion dysregulation and self-concept: an ecological momentary assessment approach. *Cogn. Emot.* 38, 1261–1270. doi: 10.1080/02699931.2024.2363413

Russ, A. L., Zillich, A. J., Melton, B. L., Russell, S. A., Chen, S., Spina, J. R., et al. (2014). Applying human factors principles to alert design increases efficiency and reduces prescribing errors in a scenario-based simulation. *J. Am. Med. Inform. Assoc.* 21, e287–e296. doi: 10.1136/amiajnl-2013-002045

Van Noordt, S. J. R., White, L. O., Wu, J., Mayes, L. C., and Crowley, M. J. (2015). Social exclusion modulates event-related frontal theta and tracks ostracism distress in children. *NeuroImage* 118, 248–255. doi: 10.1016/j.neuroimage.2015.05.085

Wang, S., and Kim, K. J. (2023). Content moderation on social media: does it matter who and why moderates hate speech? *Cyberpsychol. Behav. Soc. Netw.* 26, 527–534. doi: 10.1089/cyber.2022.0158

Wikman, P., Moisala, M., Ylinen, A., Lindblom, J., Leikas, S., Salmela-Aro, K., et al. (2022). Brain responses to peer feedback in social media are modulated by valence in late adolescence. *Front. Behav. Neurosci.* 16:790478. doi: 10.3389/fnbeh.2022.790478

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 20 October 2025
DOI 10.3389/fpsyg.2025.1699174

Check for updates

OPEN ACCESS

EDITED BY
Ivan Wen,
University of Hawaii at Manoa, United States

REVIEWED BY
Deng Cheng,
Xidian University, China
Adiyukh Berbekova,
University of Hawaii at Manoa, United States

*CORRESPONDENCE
Yuping Zhao
✉ zhaoyupingbupt@163.com
Ling Huang
✉ huang@ynu.edu.cn

†These authors have contributed equally to this work

RECEIVED 04 September 2025
ACCEPTED 02 October 2025
PUBLISHED 20 October 2025

CITATION
Zhao Y and Huang L (2025) Promoting teaching innovation among university teachers through AI literacy from the perspective of planned behavior: the moderating effects of three perceived supports.
Front. Psychol. 16:1699174.
doi: 10.3389/fpsyg.2025.1699174

Promoting teaching innovation among university teachers through AI literacy from the perspective of planned behavior: the moderating effects of three perceived supports

Yuping Zhao[1*†] and Ling Huang[2*†]

[1]General Education Center, Communication University of China, Beijing, China, [2]School of Business and Tourism Management, Postdoctoral Research Station in Business Administration, Yunnan University, Kunming, China

Introduction: The rapid development of artificial intelligence (AI) is transforming higher education, yet the mechanisms through which AI literacy influences teaching innovation among university teachers remain insufficiently explored.
Methods: This study, grounded in the Theory of Planned Behavior (TPB), investigates how AI literacy promotes teaching innovation via three psychological mechanisms: behavioral attitude, subjective norm, and perceived behavioral control. Additionally, the moderating effects of perceived support factors—teaching resources, peer support, and teaching autonomy—on the relationship between AI literacy and teaching innovation are considered. Empirical survey data from Chinese university teachers were used for analysis.
Results: The findings reveal that AI literacy significantly enhances teachers' behavioral attitudes, subjective norms, and perceived behavioral control, which in turn foster teaching innovation. Among these, perceived behavioral control plays the most significant role in driving innovative behavior. Moreover, teaching resources and teaching autonomy positively moderate the relationship between AI literacy and teaching innovation, while peer support only significantly influences behavioral attitudes.
Discussion: These results extend the application of the Theory of Planned Behavior by uncovering the psychological mechanisms through which AI literacy fosters teaching innovation. The study provides empirical evidence supporting AI literacy training and teacher support in higher education, with implications for fostering innovation in teaching practices.

KEYWORDS
Artificial Intelligence Literacy, teaching innovation, Theory of Planned Behavior, perceived support, university teachers

1 Introduction

The rapid development of artificial intelligence (AI) is profoundly transforming the ecosystem of higher education (Zhang, 2023). The widespread adoption of tools such as natural language processing, big data analysis, intelligent recommendations, and virtual teaching assistants is continuously reshaping university teachers' daily teaching practices (Niloy et al., 2025). AI has been integrated into course design, learning analytics, and educational assessment, further expanding into classroom interactions, personalized learning support, and academic monitoring (Hwang et al., 2020). The teaching model is gradually shifting from teacher-centered to learner-centered, a fundamental change driven by technological advancements (Sperling et al., 2024). This shift is part of a broader "educational paradigm transformation," where the focus moves toward more student-centered, personalized learning experiences, significantly improving teaching efficiency and flexibility (Wang et al., 2025). With the increasing application of AI technologies, university teachers are facing growing demands to update their skills and redefine their roles, shifting from traditional "knowledge transmitters" to "learning guides" and "innovative practitioners" (Kim, 2025).

Teaching innovation, as one of the key responsibilities of university teachers, is a core manifestation of professional development and a necessary condition for modernizing education and cultivating innovative talent in higher education (Wang et al., 2025). It involves the continuous exploration and improvement of teachers' educational philosophies, course goals, teaching methods, and assessment practices (Gerçek and Özveren, 2025). In this context, AI literacy has become a key factor for teachers to adapt to educational transformation and foster teaching innovation. AI literacy is generally defined as an individual's ability to understand, apply, and critically reflect on AI (Kelley and Wenzel, 2025; Ozudogru and Durak, 2025). For university teachers, it not only encompasses technical operations and tool applications but also includes the ability to evaluate educational values, identify potential risks, and creatively integrate AI into teaching (Ji et al., 2025).

Existing studies suggest that AI literacy directly influences classroom effectiveness, student experiences, and teachers' professional development (Liu et al., 2025). Teachers with higher AI literacy are more likely to break away from traditional models and demonstrate greater innovation in course design and educational assessment (Guan et al., 2025). Most research has focused on performance outcomes and technology adoption, emphasizing the relationship between literacy and tool usage, technology acceptance, and efficiency, but has insufficiently explored how AI literacy influences teaching innovation through psychological and behavioral mechanisms (Duong, 2025). In contrast, studies on student AI literacy are more systematic, while research on teachers is relatively scarce (Tzirides et al., 2024).

Therefore, investigating the impact of AI literacy on teaching innovation among teachers is of significant theoretical and practical importance. Although some scholars have suggested that AI literacy may promote teaching innovation, the underlying mechanisms remain unclear, and the impact of psychological and social factors lacks systematic explanation (Zhou et al., 2025). Existing studies often rely on models such as the Technology Acceptance Model (TAM) or Unified Theory of Acceptance and Use of Technology (UTAUT), focusing on adoption intentions while neglecting the psychological processes and perceived support environments in teaching practice (Yang et al., 2025). It is essential to reconsider the relationship between AI literacy and teaching innovation from the perspectives of psychology and organizational behavior.

The Theory of Planned Behavior (TPB) provides a solid psychological framework for understanding teachers' innovative behaviors (Zhang, 2025). According to this theory, an individual's behavioral intentions are primarily determined by behavioral attitude, subjective norm, and perceived behavioral control (Dunn et al., 2018). AI literacy may influence teachers' attitudes toward the educational value of AI, their perception of external expectations, and their confidence in their abilities, thereby promoting teaching innovation (Kong et al., 2024). At the same time, external supportive conditions cannot be overlooked in this process (Adabor et al., 2025). Perceived support theory posits that educational resources, peer support, and teaching autonomy enhance motivation and foster creative behaviors (Han et al., 2021). Educational resources provide material support for teachers to explore new methods, peer support stimulates motivation through collaborative exchange, and teaching autonomy creates institutional space for trying innovations (Cai and Tang, 2022; Hornstra et al., 2021; Nshimiyimana and Cartledge, 2020). Research has shown that a supportive environment can amplify the positive psychological and behavioral effects of an individual's capabilities (Wang et al., 2025). Thus, external situational support may moderate the impact of AI literacy on teachers' attitudes, subjective norms, and perceived control, indirectly influencing their level of teaching innovation.

In summary, existing research has the following limitations: First, there is limited systematic research on the relationship between AI literacy and teaching innovation, especially empirical studies focusing on university teachers (Chou et al., 2025); second, existing studies overly rely on technology adoption frameworks, lacking a comprehensive perspective that integrates psychology and organizational behavior (Sanusi et al., 2024); third, there is insufficient research on the role of external situational factors, such as educational resources, peer support, and teaching autonomy, in influencing teaching innovation (Ding et al., 2024; Mnguni et al., 2024).

Therefore, this study focuses on Chinese university teachers and attempts to construct and validate a comprehensive model to systematically explore how AI literacy influences teaching innovation through teachers' cognitive and psychological processes. Furthermore, the study investigates the role of external factors such as teaching resources, peer support, and teaching autonomy. This study seeks to answer the following three core questions: (1) Does AI literacy significantly promote teaching innovation among university teachers? (2) What role do behavioral attitude, subjective norm, and perceived behavioral control play in this process? (3) Does the supportive environment strengthen or weaken the relationship between AI literacy and teaching innovation?

By answering these questions, this study will not only contribute to a deeper understanding of the relationship between AI literacy and teaching innovation among university teachers but also provide empirical support and practical insights for the digital transformation of education and the professional development of teachers. The subsequent sections of the paper will present the theoretical foundation, literature review, research model and hypotheses, research methods, data analysis and results, and discussion and conclusion.

2 Theoretical framework and literature review

2.1 Theory of Planned Behavior

The Theory of Planned Behavior (TPB) posits that an individual's behavioral intentions are primarily determined by three psychological factors: behavioral attitude, subjective norm, and perceived behavioral control (Ajzen, 1991). Behavioral attitude refers to an individual's positive or negative evaluation of the likely outcomes of a specific behavior (Ahadzadeh et al., 2024). Subjective norm reflects an individual's perception of the expectations and social pressures from others in a given social context (Thanki et al., 2024). Perceived behavioral control represents an individual's assessment of their resources and abilities; the more an individual believes they have the necessary conditions and fewer potential barriers, the stronger their behavioral intention will be (Ateş, 2020). Perceived behavioral control is an internal psychological factor that determines a person's belief in their ability to perform a behavior successfully. It refers to teachers'self-efficacy, or their confidence in overcoming challenges and utilizing their abilities to incorporate AI tools into their teaching practices (Hamm et al., 2024). These three factors interact and collectively explain the formation of behavioral intentions and their translation into actual behaviors (Hou et al., 2022).

In educational research, TPB has been widely applied, particularly in explaining teachers' teaching behaviors and technology adoption (Frawley and Campbell, 2025). Studies have shown that positive behavioral attitudes enhance teachers' willingness to engage in curriculum reform and adopt new tools. External expectations and pressures, such as school policies, peer support, and student feedback, influence teaching choices through subjective norms. Teachers' confidence in their abilities and external conditions, known as perceived behavioral control, ultimately determines whether behavioral intentions translate into actual actions (Andersen et al., 2019; Gold et al., 2024).

As AI gradually integrates into higher education, teachers' AI literacy may influence all three dimensions of TPB, shaping their attitudes toward the educational value of AI, enhancing their sensitivity to social norms, and strengthening their self-efficacy (Ahadzadeh et al., 2024). Therefore, TPB provides an important theoretical framework for understanding how university teachers can achieve teaching innovation through AI-driven processes.

2.2 AI literacy

AI literacy is initially defined as an individual's ability to understand, use, and evaluate AI systems (Almatrafi et al., 2024). Its core includes not only knowledge of AI principles and mechanisms but also the ability to use AI tools effectively in real-world contexts and critically reflect on their social, ethical, and educational impacts (Ng et al., 2021). Compared to information and digital literacy, AI literacy places greater emphasis on algorithmic thinking and human-machine collaboration, and is regarded as an interdisciplinary and cross-contextual competency (Senoner et al., 2024).

As AI becomes more deeply applied in education, the concept of AI literacy has evolved from early tool-based operation to a broader competency that includes technological integration, interdisciplinary collaboration, ethical judgment, and social responsibility (Sperling et al., 2024). For teachers, AI literacy is both a prerequisite for the digital transformation of education and a critical driver of teaching innovation (Ozudogru and Durak, 2025). High levels of AI literacy can not only help teachers develop positive attitudes toward technology adoption but also reduce anxiety caused by technological uncertainty, enabling greater flexibility and creativity in course design and classroom management (Hwang and Wu, 2025).

Existing studies have identified the multidimensional characteristics of AI literacy. One stream of research emphasizes its ethical and critical dimensions, suggesting that individuals should be able to assess AI outputs and potential risks in different contexts (Kelley and Wenzel, 2025; Ozudogru and Durak, 2025). Another stream highlights the interactive dimension, noting that AI literacy involves not only cognitive and operational skills but also the ability to interact and collaborate with intelligent systems (Ayanwale et al., 2024). In the professional development of teachers, AI literacy integrates technical knowledge, teaching skills, and ethical judgment to support teachers in making informed decisions in complex educational settings (Al-Mughairi and Bhaskar, 2024; Ning et al., 2025). Overall, AI literacy for teachers is defined as a systemic competency that encompasses technical operation, algorithmic thinking, interdisciplinary integration, and social impact assessment (Abulibdeh et al., 2024; Bewersdorff et al., 2025).

Empirical studies have further validated the relationship between AI literacy and teaching innovation. Research indicates that AI literacy can not only directly promote innovative practices by enhancing teachers' technical proficiency but also indirectly foster innovation by shaping positive cognitive attitudes, strengthening the perception of social expectations, and boosting self-efficacy (Ivanov et al., 2024). Table 1 systematically reviews the latest research on AI and digital technologies in teaching innovation, providing a solid foundation for the construction of this study's model. Building on this, Figure 1 presents the evolution and application framework of AI literacy: its core is composed of the initial definitions (understanding, application, evaluation), with extensions to dimensions such as technological integration, interdisciplinary collaboration, ethical judgment, and social responsibility. These literacy components influence teaching innovation through the three psychological mechanisms of TPB—attitudes, subjective norms, and perceived behavioral control—forming a logical chain of "AI Literacy—TPB Psychological Mechanisms—Teaching Innovation" that lays a systematic foundation for the theoretical model of this study.

2.3 Teaching innovation

Teaching innovation is generally understood as the process through which teachers introduce new ideas and tools into their teaching philosophies, methods, and practices to improve learning outcomes and the teaching environment (Gilbert et al., 2021). It involves not only the adoption and integration of classroom technologies but also the transformation of course design, assessment methods, and teacher-student interaction patterns (López et al., 2023). In higher education, teaching innovation is characterized by the selection of diverse methods, integration of interdisciplinary resources, and personalized responses to learners' needs, making teaching more flexible, open, and adaptive (Miranda et al., 20214).

Existing research generally agrees that teaching innovation is influenced by both individual and contextual factors (Kottmann et al., 2024). On the individual level, a teacher's knowledge structure, innovation awareness, and technical abilities determine the likelihood of implementing changes in their teaching (Chen, 2024). On the contextual level, institutional support, peer collaboration, and technological environments have been identified as key conditions for promoting teaching innovation (Mokhlis and Abdullah, 2024). In recent years, the widespread application of AI has further expanded the boundaries of teaching innovation. It not only provides tools like learning analytics and intelligent feedback but also facilitates the paradigm shift from a "teacher-centered" to a "learner-centered" model (Chou et al., 2025). However, teachers still face challenges in

TABLE 1 Research progress on AI and digital technologies in teaching innovation.

Reference	Research context	Research method	Research finding
Panday-Shukla (2025)	AI literacy	Quantitative Research	Pre-service teachers and teacher educators have moderate digital literacy but low AI literacy.
Ozudogru and Durak (2025)	Artificial Intelligence	Quantitative Research	AI readiness (cognition, vision, and ethics) significantly impacts AI-enhanced innovation levels in teaching.
Chen et al. (2025)	AI Technologies	Quantitative Research	Adequate technical support and adaptable AI tools are crucial for integrating AI into STEM education.
Chu and Wang (2025)	AI-Integrated	Quantitative Research	Micro and individual factors, especially beliefs in AI's potential, significantly impact teachers' epistemic agency, fostering innovation.
Chen and Zou (2024)	Intelligent MR devices	Quantitative Research	Intelligent teaching devices enhance educational equity and teaching quality, particularly in remote areas.
Robayo-Pinzon et al. (2024)	Artificial Intelligence	Quantitative Research	Students generally agree with the co-creation of value through AI functions in higher education scenarios.
Lin and Chen (2024)	Artificial intelligence	Quantitative Research	AI applications can constrain creativity and innovation due to rigid frameworks.
Kim (2024)	Artificial intelligence	Quantitative Research	Optimizing the complementary strengths of both human teachers and AI holds great potential for educational innovation.

Source(s): created by author.

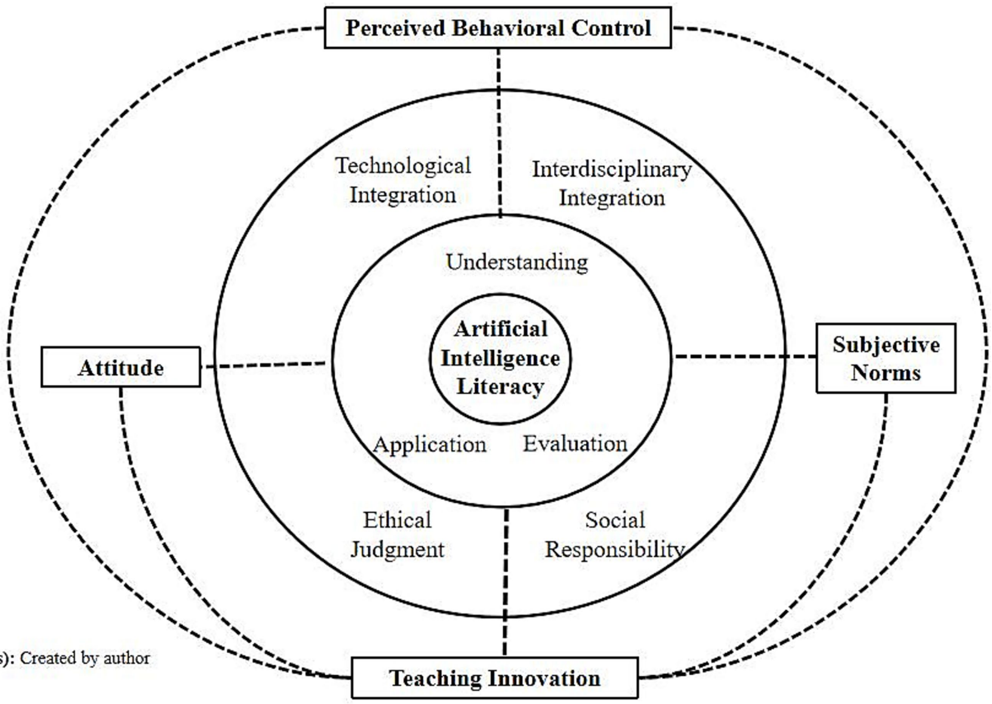


FIGURE 1
Evolution of AI literacy and the framework for teacher teaching innovation.

advancing innovation, including insufficient AI literacy, limited teaching autonomy, and uncertainty about new teaching models (Gupta and Bhaskar, 2020). Table 2 summarizes the latest research on teaching innovation, covering the research context, methods, and key findings. These results indicate that teacher innovation relies both on individual cognition and attitudes and on the important influence of external support environments. This lays the foundation for exploring the "AI literacy—TPB psychological mechanisms—teaching innovation" pathway in this study.

2.4 Perceived support

Perceived support refers to an individual's subjective perception of the available resources, social relationships, and autonomy within an organizational context. It is widely recognized as a critical psychosocial factor influencing motivation, behavior, and innovation (Wahid and Ayub, 2024). In the context of higher education, perceived support for teachers not only stems from institutional guarantees and material resources but also includes emotional recognition and social

TABLE 2 Overview of research progress on teaching innovation.

Reference	Research context	Research method	Research finding
Xiang et al. (2024)	Career calling	Quantitative Research	Career calling is positively correlated with teacher innovation.
Cai and Tang (2021)	School support	Quantitative Research	The impact of school support for innovation on teacher innovation varies.
Liu et al. (2024)	Conceptualizations	Qualitative Research	Domain-specific definitions aid in understanding teacher innovation.
Liu et al. (2022)	Professional learning communities	Quantitative Research	School-level professional learning communities positively influence individual teacher innovation.
Han et al. (2021)	Perceived support	Qualitative Research	The relationship between teaching resources and teacher innovation is minimal.
Teng et al. (2024)	Distributed leadership	Qualitative Research	Distributed leadership impacts teacher innovation at both team and individual levels.
Ertas and Pekmezci (2025)	Career motivation	Qualitative Research	Instructional practice and teacher innovation mediate the relationship between social utility motivation and job satisfaction.
Bao (2024)	Principals' secure base leadership	Qualitative Research	Principals' secure base leadership enhances teacher innovation through affective commitment.
Qin et al. (2025)	Teacher collaboration	Qualitative Research	Teacher collaboration significantly boosts innovation ability and teaching motivation.
Ma and Zhang (2025)	Distributed leadership	Qualitative Research	Distributed leadership does not directly predict teacher innovation behavior.
Adams et al. (2025)	Openness to experience	Qualitative Research	Teachers' openness to experience significantly predicts creativity, LMX quality, and innovative teaching practices.

Source(s): created by author.

support derived from peer collaboration, organizational atmosphere, and management mechanisms (Cai and Tang, 2022). These elements help to stimulate positive teaching attitudes and enhance innovative motivation (Liu and Chang, 2023).

In the context of university teachers adapting to and applying AI technologies in their teaching practices, perceived support can be broken down into three key dimensions: teaching resources, peer support, and teaching autonomy (Kruse et al., 2024). These three dimensions are distinct but interrelated, and together they provide a comprehensive framework of support that enables teachers to navigate the challenges of AI integration in teaching.

Teaching resources, such as AI training opportunities, access to digital tools, and platform infrastructure, provides the technological foundation for teachers to improve their AI literacy and drive classroom innovation (Padilha et al., 2021). High levels of resource support enhance teachers' understanding and control over AI tools, lowering the barriers to technology adoption and increasing their willingness to actively incorporate AI into their teaching (Cai and Tang, 2022).

Peer support plays a buffering and motivating role in the adoption of AI technologies (Hornstra et al., 2021). Collaboration and communication among teachers not only help share experiences of AI teaching practices and reduce the uncertainty associated with technology, but also provide emotional support and a sense of belonging, thereby boosting teachers' technological confidence and innovative motivation (Adie et al., 2024). This is especially important in the context of rapid AI tool iterations.

Teaching autonomy refers to the freedom and decision-making power that teachers have in course design, teaching methods, and the selection of teaching tools (Zhao and Qin, 2021). Teachers with higher levels of teaching autonomy are better able to independently adjust and innovate their teaching methods, particularly as AI tools become integrated into their teaching practices (Martinek et al., 2020). Teaching autonomy enhances teachers' sense of ownership over AI integration, enabling them to adapt AI tools to better meet the needs of their students and teaching objectives (Vangrieken and Kyndt, 2020). It facilitates the transformation of technical competence into classroom practices and encourages teachers to adopt new, creative approaches in response to the dynamic educational landscape (Bali et al., 2025).

As AI continues to be embedded in educational practices, the diversity of educational resources, the continuity of peer support, and the enhancement of teaching autonomy provide critical psychological foundations for teachers to translate AI literacy into teaching innovation behaviors (Okada, 2023; Zhao et al., 2021). These three types of perceived support not only mitigate psychological barriers during the technology adoption process but also stimulate teachers' sense of technological efficacy and autonomy, playing an irreplaceable role in moderating and empowering the integration of AI in teaching.

3 Research model and hypotheses development

3.1 AI literacy and the Theory of Planned Behavior (behavioral attitude, subjective norm, and perceived behavioral control)

AI literacy reflects a teacher's comprehensive understanding of knowledge, operational skills, and critical thinking, which influences their acceptance and use of AI tools in teaching contexts (Ma and Lei, 2024). Teachers with higher AI literacy are more likely to recognize the potential of AI in enhancing classroom efficiency and improving learning experiences, gradually forming a positive attitude (Wang and Wang, 2024). Attitude is not simply an emotional preference but represents a deeply cognitive and value-based stance toward AI integration. Teachers with positive attitudes are more likely to engage with AI technologies, incorporating them into course design, classroom interactions, and assessment methods (To et al., 2023). This stable orientation provides the psychological momentum for teaching innovation, making innovative behaviors more common. AI literacy also influences teachers' perceptions of external norms. In line with the Theory of Planned Behavior (TPB), teachers' behaviors are significantly influenced by the subjective norms around them, such as policy support, disciplinary communities, and student expectations (Dierendonck et al., 2024). Teachers with higher AI literacy are more likely to internalize these external norms as part of their professional identity, strengthening their sense of responsibility and enhancing their innovative behavior (Adelana et al., 2024). Perceived behavioral control is similarly affected by AI literacy. AI literacy enhances teachers' sense of self-efficacy, allowing them to manage classroom uncertainty, break down tasks, and remain confident even in the face of technical difficulties (Chen et al., 2023). This is particularly crucial under limited resources, as it reduces psychological resistance caused by uncertainty and increases teachers' willingness to engage in innovative behaviors (Hamm et al., 2024). AI literacy shapes teachers' psychological readiness and behavioral tendency for teaching innovation through the three dimensions of attitude, subjective norm, and perceived behavioral control. Based on this, the following hypotheses are proposed:

> *H1a. AI literacy has a significant positive effect on teachers' behavioral attitude.*
>
> *H1b. AI literacy has a significant positive effect on teachers' subjective norm.*
>
> *H1c. AI literacy has a significant positive effect on teachers' perceived behavioral control.*

3.2 The Theory of Planned Behavior (behavioral attitude, subjective norm, and perceived behavioral control) and teaching innovation

In the teaching domain, the three components of the Theory of Planned Behavior form the key psychological foundation for teachers' innovation. Behavioral attitude represents teachers' cognition and emotional experience regarding the value of AI in teaching (Liu and Wang, 2024). Positive attitudes, informed by cognitive appraisals and emotional investments, are critical in fostering teachers' willingness to experiment with new tools, restructure course plans, and engage in repeated trials, all of which enhance the scope and depth of innovation. As the attitude becomes more stable, teachers are more inclined to adopt structured and adaptive methods in course design, incorporating intelligent feedback, layered support, and data-driven evaluation, thus extending innovative practices (Ehlert et al., 2022). Subjective norm represents the societal and professional expectations placed on teachers. This dimension underscores the influence of external norms, such as policy guidelines, peer practices, and student demands, in shaping teachers' professional responsibilities (Knauder and Koschmieder, 2019). Teachers' internalization of these norms not only strengthens their social responsibility but also increases their commitment to adopting innovations, as external pressures and professional values align. Perceived behavioral control reflects teachers' judgment of feasibility and control during the innovation process. Teachers with stronger control can break down complex goals into manageable tasks, maintain steady progress in resource-constrained situations, and use data feedback for continuous improvement (Zhan et al., 2024). Perceived behavioral control, as influenced by self-efficacy, determines how confidently teachers can face challenges, overcome failures, and persist in innovative efforts, transforming the innovation process from trial and error to a sustained, systematic practice. Attitude, norm, and control impact cognition, social aspects, and operations, respectively, collectively driving teachers to transform potential intentions into visible practices, forming an intrinsic motivation system for teaching innovation (Ateş, 2020). Based on this, the following hypotheses are proposed:

> *H2a. Behavioral attitude has a significant positive effect on teachers' teaching innovation.*
>
> *H2b. Subjective norm has a significant positive effect on teachers' teaching innovation.*
>
> *H2c. Perceived behavioral control has a significant positive effect on teachers' teaching innovation.*

3.3 The moderating role of perceived support (teaching resources, peer support, teaching autonomy)

Teaching resources are key conditions for teachers to engage in innovative practices, encompassing hardware, software platforms, training opportunities, and institutional support (Wu et al., 2022). According to the Theory of Planned Behavior, environmental conditions significantly influence the formation of attitudes, norms, and perceived control (Wang et al., 2025). The availability of resources determines whether teachers can effectively translate their AI literacy into positive psychological mechanisms (Ateş, 2020). In teaching contexts, abundant resources provide both material and emotional support, helping to build teachers' confidence in the application of AI tools. Regarding attitudes, abundant resources reduce the risks of practice, making it easier for teachers to translate literacy into positive evaluations. Equipment and services provide a safety net, creating value convictions and emotional investment during operations (Ayanwale et al., 2025). This material and emotional safety net helps solidify teachers' commitment to AI integration and teaching innovation, reducing psychological barriers to innovation. Subjective norms also depend on resource support. Resources not only provide material conditions but also symbolize the organization's and community's focus on AI teaching, leading teachers to perceive stronger external recognition and expectations (Ramnarain et al., 2024). As resources grow, teachers perceive a stronger alignment with institutional and professional goals, reinforcing their commitment to innovation. Perceived behavioral control is more closely related to resources. Available tools and services give teachers more control in complex situations, enhancing self-efficacy and promoting the realization of innovation intentions (Gong, 2023). Resources act as "magnifiers." While literacy provides knowledge and skills, the positive effects of literacy are hard to fully utilize in the absence of resources. When resources are sufficient, the positive effects of literacy on attitude, norms, and control are strengthened, making it easier for innovation motivation to be converted into action. The following hypotheses are proposed:

> *H3a. Teaching resources positively moderate the relationship between AI literacy and behavioral attitude.*
>
> *H3b. Teaching resources positively moderate the relationship between AI literacy and subjective norm.*
>
> *H3c. Teaching resources positively moderate the relationship between AI literacy and perceived behavioral control.*

Peer support reflects the emotional encouragement, experience sharing, and role modeling teachers receive within teams and academic communities. Social support theory indicates that positive peer interactions can alleviate stress and enhance innovation confidence (Wu et al., 2022). From the perspective of the Theory of Planned Behavior, peer support, as an important aspect of the social environment, has a significant influence on the formation of attitudes, subjective norms, and perceived behavioral control (Frawley and Campbell, 2025). Peer support, through collaborative interactions and shared experiences, reduces the isolation teachers may face and amplifies the social and emotional aspects of innovation. On the attitude level, even if teachers possess AI literacy, without peer encouragement, it is difficult to transform cognitive advantages into emotional investment (Zhou et al., 2022). A positive team atmosphere and practical demonstrations help build confidence and positive emotions (Sokha, 2024). Subjective norms are strengthened by peer support. Compared to policy documents, the adoption and demonstration by colleagues are more persuasive, leading teachers to perceive group recognition and internalize it as professional responsibility (Zhao et al., 2024). The collective validation from peers helps solidify teachers' understanding of their innovation efforts as valid and valuable within their professional community. Perceived behavioral control also benefits from peer support. Collaboration and mutual assistance prevent teachers from facing technical or teaching challenges in isolation, enhancing control and willingness to act (Wan et al., 2024). Peer support not only shares resources but also provides social validation. Teachers, in a group-acknowledged environment, feel the practical value of their efforts and are more likely to transform innovation into normalized behavior. Therefore, peer support can strengthen the effect of AI literacy on the elements of the Theory of Planned Behavior, turning it into genuine innovative motivation. The following hypotheses are proposed:

> *H4a. Peer support positively moderates the relationship between AI literacy and behavioral attitude.*
>
> *H4b. Peer support positively moderates the relationship between AI literacy and subjective norm.*
>
> *H4c. Peer support positively moderates the relationship between AI literacy and perceived behavioral control.*

Teaching autonomy reflects the degree of freedom teachers have in course design, teaching methods, and tool selection. Self-determination theory emphasizes that autonomy can stimulate intrinsic motivation, increasing engagement and innovation willingness (Reeve and Cheon, 2024). In the framework of the Theory of Planned Behavior, autonomy is an important external condition affecting attitudes, norms, and perceived control, determining whether AI literacy can translate into positive psychological mechanisms (Ren, 2024). Autonomy provides teachers with a sense of ownership over their teaching, which in turn enhances the value they place on innovation and the integration of AI. On the attitude level, teachers with AI literacy, but limited in teaching activities, find it difficult to form positive emotions. As autonomy increases, teachers can freely apply AI tools based on their preferences, creating value convictions (Vangrieken and Kyndt, 2020). Subjective norms are more likely to internalize due to autonomy. Teachers can combine external requirements with personal will, shifting from passive compliance to professional recognition (Martinek et al., 2020). Perceived behavioral control also depends on autonomy. Greater freedom reduces external barriers, enhancing teachers' sense of control and self-efficacy (Miao and Ma, 2023). By increasing control over their teaching practices, autonomy allows teachers to overcome external challenges and strengthens their commitment to innovation. Autonomy enhances confidence and reduces resistance, making it an essential condition for transforming innovation intentions into practice. Teachers in an autonomous environment are more likely to explore and gradually form stable innovation patterns. Therefore, teaching autonomy not only directly promotes teaching innovation but also strengthens the effect of AI literacy on the elements of the Theory of Planned Behavior, making psychological motivation more likely to turn into action. The following hypotheses are proposed:

> *H5a. Teaching autonomy positively moderates the relationship between AI literacy and behavioral attitude.*
>
> *H5b. Teaching autonomy positively moderates the relationship between AI literacy and subjective norm.*
>
> *H5c. Teaching autonomy positively moderates the relationship between AI literacy and perceived behavioral control.*

3.4 Mediating role of the Theory of Planned Behavior

AI literacy not only directly affects teachers' teaching innovation but also exerts an indirect effect through the three core components of the Theory of Planned Behavior: behavioral attitude, subjective norm, and perceived behavioral control (Ma, 2025). These three dimensions constitute the psychological mechanisms that enable teachers' knowledge and skills to be transformed into visible innovative behaviors. In the behavioral attitude dimension, higher AI literacy helps teachers understand the value of AI in enhancing classroom efficiency, improving learning experiences, and achieving personalized support (Liu and Wang, 2024). Recognition of these values gradually accumulates into positive emotional experiences and solidifies into a positive attitude toward AI applications. Positive attitudes guide teachers to more readily experiment with tools, adjust processes, and conduct small-scale experiments in

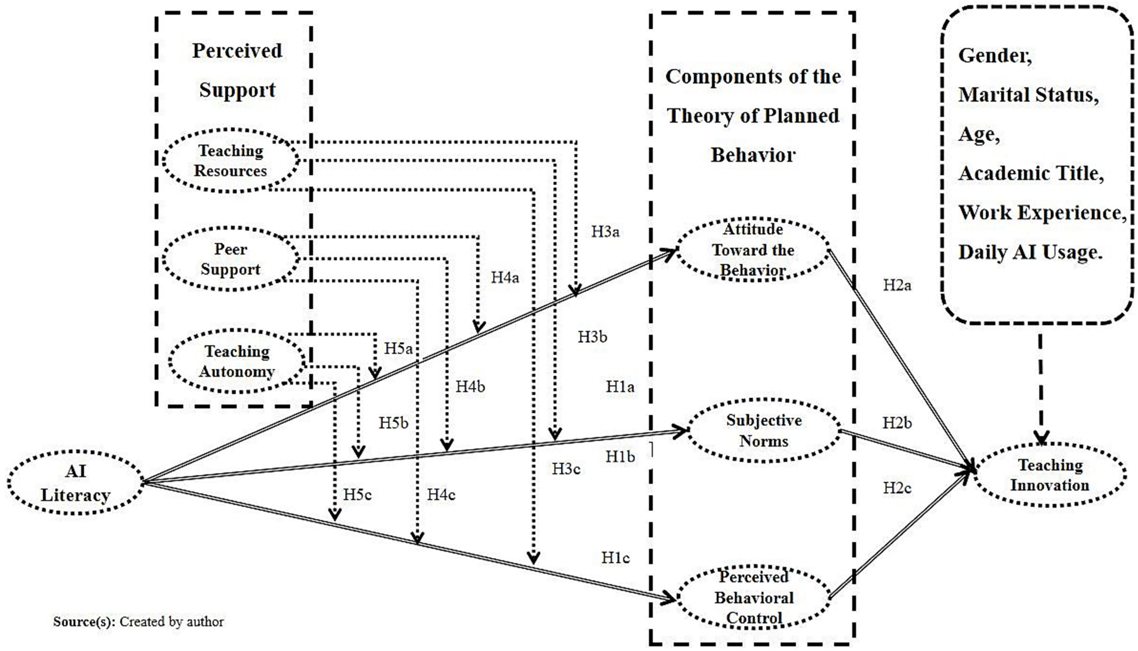


FIGURE 2
Research model.

teaching practice, thus enhancing the continuity and scope of innovation activities. In the subjective norm dimension, AI literacy increases teachers' sensitivity to the external environment (Adelana et al., 2024). Teachers can accurately interpret policy directions, peer practices, and student needs, perceiving widespread recognition of teaching innovation within the professional community. This recognition reinforces teachers' social responsibility, transforming external pressure into self-identity, making innovation a natural choice for teaching. In the perceived behavioral control dimension, AI literacy strengthens teachers' tool usage and problem-solving abilities, enhancing self-efficacy (Lim, 2023). Teachers believe they have the ability to deal with technical problems, classroom uncertainty, and resource shortages. Control enhances teachers' confidence and stability when facing challenges, making innovative activities no longer high-risk trials, but sustainable routine practices. Attitude, subjective norm, and perceived behavioral control form the key psychological path through which AI literacy affects teaching innovation. These three factors work together, enabling teachers to move from having the capability to willingness to action, and ultimately to sustained innovation. The following hypotheses are proposed:

> *H6a. Behavioral attitude mediates the relationship between AI literacy and teaching innovation.*
>
> *H6b. Subjective norm mediates the relationship between AI literacy and teaching innovation.*
>
> *H6c. Perceived behavioral control mediates the relationship between AI literacy and teaching innovation.*

In summary, the research model of this study is shown in Figure 2.

4 Method

4.1 Participants

The data for this study were collected online via the Wenjuanxing platform,[1] using a seven-point Likert scale. To ensure sample relevance, two mandatory screening questions were placed at the start: (1) Occupational identity (only university teachers); (2) Experience using AI tools (must answer "Yes"). Respondents failing either screen were blocked from proceeding to the main questionnaire. Data cleaning followed a pre-specified process to ensure quality and consistency: (1) Duplicate response detection was performed, and surveys with duplicate responses from the same IP address or device were checked and excluded to avoid redundancy; (2) Response consistency was examined, and surveys where all items were answered with the same option were excluded; (3) Incomplete responses were removed. In total, 518 questionnaires were collected. After applying these criteria, 15 responses were excluded, resulting in 503 valid responses (effective response rate 97.1%). The valid sample represents a variety of universities across the country. The gender distribution was fairly balanced, with the majority of participants aged between 25 and 46 years. The sample included assistant professors, lecturers, associate professors, and professors, with years of experience ranging from 0–5 years to over 20 years. Most teachers reported using AI tools for more than 1 h daily, with some using them for over 5 h. Common platforms included ChatGPT, DeepSeek, Sora, and Wenxin Yiyan, indicating the widespread integration of AI in teaching and research practices. The demographic characteristics of the participants are shown in Table 3.

1 https://www.wjx.cn/vm/Y8VN9Xp.aspx#

TABLE 3 Demographic characteristics.

Variable	Category	Frequency	Percentage
Gender	Male	274	54.47%
	Female	229	45.53%
Marital status	Married	349	69.38%
	Unmarried	154	30.62%
Age	25–35	215	42.74%
	36–46	156	31.01%
	47–57	96	19.09%
	58 and above	36	7.16%
Academic title	Teaching Assistant/No Title	194	38.57%
	Lecturer	178	35.39%
	Associate Professor	97	19.28%
	Professor	34	6.76%
Work experience	0–5 years	88	17.5%
	6–10 years	167	33.2%
	11–15 years	105	20.87%
	16–20 years	96	19.09%
	More than 21 years	47	9.34%
Daily AI usage	0–1 h	28	5.57%
	1–3 h	134	26.64%
	3–5 h	141	28.03%
	5–7 h	106	21.07%
	More than 7 h	94	18.69%

Source(s): created by author.

4.2 Measures

This study utilized established scales that have been empirically validated both domestically and internationally, with modifications made to suit the context of university teachers (Cui and Yin, 2023; Han et al., 2021; Liao et al., 2022; Liu et al., 2016; Ning et al., 2025; Richter and Schuessler, 2019; Wang et al., 2023; Zhang et al., 2025). All items were rated using a seven-point Likert scale (1 = strongly disagree, 7 = strongly agree), with higher scores indicating higher levels on each dimension. The measurement covered eight core variables: AI literacy (9 items, assessing teachers' abilities to recognize, apply, evaluate AI tools, and their awareness of ethical risks), teaching innovation (5 items, reflecting practices such as exploring new ideas, applying diverse teaching methods, problem-solving, sharing experiences, and integrating resources), behavioral attitude (3 items, measuring teachers' positive cognitive responses to organized teaching and research activities), subjective norm (3 items, reflecting social expectations from the department, colleagues, and academic groups), perceived behavioral control (3 items, addressing factors such as time, channels, and self-efficacy), teaching resources (3 items, evaluating institutional support for training, tools, and hardware facilities), peer support (3 items, reflecting experience sharing, encouragement, and collaboration among colleagues), and teaching autonomy (3 items, reflecting teachers' freedom in decision-making related to the integration of AI in teaching). The specific items and scale sources for each variable are listed in Table 4.

4.3 Common method Bias analysis

Since this study used self-reported questionnaires to collect data, there is a potential risk of common method bias (Podsakoff et al., 2003). To minimize this issue, several measures were implemented during the questionnaire design, including ensuring anonymity, adjusting the order of items, and incorporating attention check questions. In terms of statistical testing, Harman's single-factor analysis was conducted. The results indicated that the first factor explained 28.213% of the variance, which is well below the 40% threshold, suggesting that a single factor did not dominate the data. Additionally, the variance inflation factors (VIFs) for the latent variables were all below 3.3 (Kock, 2015), further indicating that common method bias poses a limited threat to the study's results.

5 Data analysis and results

This study employed Partial Least Squares Structural Equation Modeling (PLS-SEM) as the primary analytical method. PLS-SEM is suitable for complex models involving multiple latent variables and interaction effects, providing robust estimation results for both measurement and structural models (Henseler et al., 2015). PLS-SEM requires fewer assumptions regarding data distribution, making it particularly appropriate for exploratory and prediction-oriented research (Sarstedt et al., 2022). Compared to traditional covariance-based structural equation models, PLS-SEM offers greater flexibility and predictive power, especially in path analysis with multiple latent variables and indicators (Rigdon, 2016). Considering the inclusion of multiple core variables such as AI literacy, perceived support, teaching attitude, and teaching innovation, as well as the need to test their complex relationships, PLS-SEM aligns well with the research requirements and methodological approach (Carrión et al., 2017).

5.1 Measurement model evaluation

To assess the reliability and validity of the measurement tools, internal consistency and convergent validity of the latent variables were first analyzed (Table 5). The results showed that the Cronbach's α values for all latent variables were above 0.80, and composite reliability (CR) exceeded 0.70, indicating strong internal consistency. Additionally, the standardized factor loadings for all measurement items were greater than 0.70 and significant, with average variance extracted (AVE) exceeding 0.50, meeting the recommended standards by Cheung et al. (2024), indicating good convergent validity for the constructs. In terms of discriminant validity, heterotrait-monotrait (HTMT) ratio analysis was conducted (Table 6). The results showed that the HTMT values for all variable pairs were below 0.85, consistent with the threshold set by Henseler et al. (2015), indicating good discriminant validity among the latent variables. Overall, the

TABLE 4 Measurement scales.

Variable	Item	Item description	Scale source
Artificial Intelligence Literacy (AIL)	AIL1	I can distinguish between AI-powered and non-AI-powered devices.	Ning et al. (2025), Wang et al. (2023)
	AIL2	I can identify AI technologies used in the applications or products I use daily.	
	AIL3	I understand how to apply AI tools to improve my teaching or research efficiency.	
	AIL4	I am proficient in using AI-related applications or products for teaching or research tasks.	
	AIL5	I can select the most appropriate AI tool or platform based on specific task requirements.	
	AIL6	I can assess the strengths and limitations of AI applications.	
	AIL7	When presented with multiple suggestions from an intelligent system, I can choose the most suitable solution.	
	AIL8	I actively consider ethical and privacy issues when using AI tools.	
	AIL9	I remain vigilant about the potential misuse of AI technologies in teaching or research.	
Teaching Innovation (TI)	TI1	I actively explore and experiment with new teaching concepts to enhance students' cognitive engagement.	Cui and Yin (2023), Liu et al. (2016)
	TI2	I regularly apply diverse teaching methods or technologies in class to stimulate students' interest in learning.	
	TI3	When faced with teaching challenges, I proactively adopt new strategies or approaches to solve problems.	
	TI4	I am willing to share my experiences of implementing new teaching ideas or methods with colleagues to receive feedback and support.	
	TI5	To achieve teaching innovation, I actively seek out and integrate necessary resources and tools (such as AI technologies, ICT, etc.).	
Behavior Attitude (BA)	BA1	I believe participating in organized research activities helps me gain more knowledge and academic resources in my field.	Liao et al. (2022), Zhang et al. (2025)
	BA2	I believe organized research activities improve my research efficiency.	
	BA3	I believe participating in organized research activities enhances the quantity and quality of my research output.	
Subjective Norm (SN)	SN1	My school or department encourages faculty to participate in organized research activities.	
	SN2	I believe my colleagues, mentors, or supervisors expect me to actively engage in organized research activities.	
	SN3	Many young faculty members around me are beginning to participate in organized research activities.	
Perceived Behavioral Control (PBC)	PBC1	I have sufficient time to participate in organized research activities.	
	PBC2	I am aware of the channels or platforms through which I can participate in organized research.	
	PBC3	I believe I have the necessary skills and experience to engage in organized research activities.	

(Continued)

10.3389/fpsyg.2025.1699174

TABLE 4 (Continued)

Variable	Item	Item description	Scale source
Teaching Resources (TR)	TR1	My institution provides professional training and guidance on integrating AI technologies into teaching.	Han et al. (2021), Richter and Schuessler (2019)
	TR2	I have access to AI-related software and technological tools provided by my institution.	
	TR3	The school is equipped with basic hardware infrastructure that supports AI-integrated teaching (e.g., smart classrooms, teaching terminals).	
Peer Support (PR)	PR1	My colleagues are willing to offer advice and share their experiences on using AI in teaching.	
	PR2	When I encounter difficulties using AI technologies in teaching, I can receive encouragement and support from my colleagues.	
	PR3	My colleagues actively support and explore ways to integrate AI technologies into teaching.	
Teaching Autonomy (TA)	TA1	I have the autonomy to decide whether to integrate AI technologies into my teaching.	
	TA2	I can independently choose appropriate AI tools and methods based on my teaching goals.	
	TA3	I have the freedom to flexibly incorporate AI technologies throughout the teaching process.	

Source(s): created by author.

measurement model achieved high levels of reliability, convergent validity, and discriminant validity.

5.2 Structural model analysis

To control for potential confounding effects, demographic variables such as gender, marital status, age, academic title, work experience, and AI usage duration were included as control variables (Cui and Yin, 2023; Jose et al., 2025; Zhao et al., 2022). The analysis indicated that these control variables did not have a significant impact on the main relationships, specifically the relationships between AI literacy, the components of the Theory of Planned Behavior (behavioral attitude, subjective norm, and perceived behavioral control), and teaching innovation. This suggests that the main relationships between AI literacy and teaching innovation are not influenced by these demographic factors, and the results are consistent across different population subgroups.

5.2.1 Model explained variance and predictive relevance

Table 7 displays the explanatory power and predictive relevance of the structural model. The results indicated that the R^2 values for the endogenous variables ranged from 0.269 to 0.309, indicating that the exogenous variables explained a substantial portion of the variance in the endogenous constructs. Additionally, the Stone-Geisser Q^2 values for all endogenous variables were greater than zero (ranging from 0.190 to 0.224), suggesting strong robustness and reliability of the model in out-of-sample predictions. Overall, these results support the model's rationality from both explanatory power and predictive relevance perspectives, further highlighting its theoretical value (Sarstedt et al., 2020; Shmueli et al., 2019).

5.2.2 Main effects path coefficient analysis and hypothesis testing

Figure 3 presents the results of the main effects path coefficient tests for the structural model. AI literacy has a significant positive effect on teachers' behavioral attitude ($\beta = 0.159$), subjective norm ($\beta = 0.224$), and perceived behavioral control ($\beta = 0.292$), supporting hypotheses H1a, H1b, and H1c. These results align with Ivanov et al. (2024), indicating that higher AI literacy helps teachers form positive teaching attitudes, enhances their perception of external normative expectations, and boosts their self-efficacy in applying AI in teaching. Furthermore, behavioral attitude ($\beta = 0.146$), subjective norm ($\beta = 0.189$), and perceived behavioral control ($\beta = 0.344$) all significantly predict teaching innovation, supporting hypotheses H2a, H2b, and H2c. Perceived behavioral control had the most significant effect, confirming Broadbent et al. (2024), Opoku et al. (2021)'s finding that teachers are more likely to engage in innovative practices when they perceive greater control over teaching.

5.2.3 Moderating effect analysis

Figure 3 illustrates the moderating effects of teaching resources, peer support, and teaching autonomy on the relationships between AI literacy and teachers' behavioral attitude, subjective norm, and perceived behavioral control. Teaching resources showed significant positive moderating effects on all three paths. As the level of teaching resources increased, the influence of AI literacy on behavioral attitude

TABLE 5 Results of reliability and convergent validity testing.

Latent variable	Measurement items	Mean	Standard deviation	Factor loading	Cronbach's α	CR	AVE
AIL	AIL1	4.507	1.539	0.773	0.920	0.932	0.606
	AIL2	4.513	1.572	0.792			
	AIL3	4.507	1.544	0.762			
	AIL4	4.473	1.554	0.787			
	AIL5	4.541	1.554	0.767			
	AIL6	4.489	1.529	0.782			
	AIL7	4.515	1.574	0.779			
	AIL8	4.513	1.520	0.788			
	AIL9	4.406	1.518	0.775			
TI	TI1	4.197	1.607	0.795	0.866	0.902	0.649
	TI2	4.181	1.601	0.793			
	TI3	4.241	1.594	0.806			
	TI4	4.203	1.510	0.811			
	TI5	4.215	1.663	0.823			
BA	BA1	4.320	1.679	0.863	0.840	0.904	0.758
	BA2	4.235	1.632	0.867			
	BA3	4.302	1.643	0.881			
SN	SN1	4.491	1.632	0.863	0.826	0.895	0.740
	SN2	4.489	1.621	0.858			
	SN3	4.491	1.619	0.859			
PBC	PBC1	4.433	1.651	0.871	0.836	0.901	0.752
	PBC2	4.408	1.689	0.852			
	PBC3	4.306	1.648	0.879			
TR	TR1	4.654	1.571	0.827			
	TR2	4.654	1.603	0.875			
	TR3	4.682	1.587	0.860			
PS	PS1	4.726	1.490	0.846	0.817	0.887	0.723
	PS2	4.791	1.474	0.824			
	PS3	4.805	1.603	0.880			
TA	TA1	4.718	1.690	0.879	0.838	0.900	0.750
	TA2	4.686	1.587	0.857			
	TA3	4.789	1.614	0.863			

Source(s): created by author.

TABLE 6 Results of discriminant validity testing.

Variable	AIL	BA	SN	PBC	TI	TR	PS	TA
AIL								
BA	0.362							
SN	0.405	0.506						
PBC	0.478	0.466	0.557					
TI	0.413	0.423	0.486	0.573				
TR	0.163	0.396	0.356	0.370	0.375			
PS	0.186	0.357	0.385	0.319	0.330	0.395		
TA	0.162	0.320	0.339	0.345	0.326	0.475	0.422	

Source(s): created by author.

($\beta = 0.101, p = 0.043$), subjective norm ($\beta = 0.112, p = 0.017$), and perceived behavioral control ($\beta = 0.107, p = 0.021$) strengthened, supporting hypotheses H3a, H3b, and H3c. This indicates that sufficient teaching resources provide perceived support for teachers and further enhance their positive attitudes and beliefs in the use of emerging technologies (Hazzan-Bishara et al., 2025).

Peer support had a significant moderating effect only on the relationship between AI literacy and behavioral attitude ($\beta = 0.170, p = 0.001$), supporting hypothesis H4a. However, the effects on subjective norm ($\beta = 0.020, p = 0.653$) and perceived behavioral control ($\beta = 0.017, p = 0.741$) were not significant, and hypotheses H4b and H4c were not supported. This suggests that peer support is context-dependent in teachers' technology adoption, more likely to influence the attitude dimension rather than universally affect all cognitive factors (Celik et al., 2025; Habibi et al., 2023).

Teaching autonomy exhibited significant positive moderating effects on all three paths. The higher the teaching autonomy, the stronger the impact of AI literacy on behavioral attitude ($\beta = 0.097, p = 0.036$), subjective norm ($\beta = 0.106, p = 0.016$), and perceived behavioral control ($\beta = 0.133, p = 0.004$), supporting hypotheses H5a, H5b, and H5c. This result emphasizes the key role of teaching autonomy in fostering technology adoption and innovation practices, indicating that empowerment and decision-making autonomy effectively stimulate teachers' proactivity and initiative in applying AI technologies (Bali et al., 2025; Hou and Shen, 2024).

Furthermore, to visually present the moderating effects, interaction effect plots for teaching resources, peer support, and teaching autonomy were generated (Figure 4), and the effects at different levels were reported (Table 8). The results showed that teaching resources had significant positive moderating effects on all three paths, with low-level effects being non-significant, medium-level effects enhancing, and high-level effects being the strongest (H3a–H3c). Peer support had significant effects only on the behavioral attitude path (H4a), with no significant effects on subjective norm or perceived behavioral control (H4b and H4c). Teaching autonomy exhibited significant positive moderating effects on all three paths, and the effects strengthened as the level of autonomy increased (H5a–H5c).

5.2.4 Mediation effect analysis

Using SmartPLS, the mediation effects of behavioral attitude (BA), subjective norm (SN), and perceived behavioral control (PBC) in the relationship between AI literacy (AIL) and teaching innovation (TI) were tested based on 5,000 bootstrap samples. According to Hayes (2009), indirect effects are considered significant if the 95% confidence interval does not include zero. The results (Table 9) indicated that all three mediation paths were significant, with no confidence intervals crossing zero. Specifically, the indirect effect of AI literacy through BA was relatively small ($\beta = 0.023, p < 0.001$), supporting H6a; the effect through SN was moderate ($\beta = 0.042, p < 0.001$), supporting H6b; and the effect through PBC was the largest ($\beta = 0.101, p < 0.001$), supporting H6c. Additionally, the direct effect of AI literacy on TI remained significant ($\beta = 0.166, p < 0.001$), indicating that BA, SN, and PBC partially mediate the relationship between AI literacy and teaching innovation. These findings confirm Ramnarain et al. (2024), suggesting that AI literacy not only directly enhances teachers' innovation inclination but also indirectly boosts innovation momentum through multiple psychological mechanisms (attitudes, norms, control beliefs).

6 Discussion

AI literacy significantly enhances teachers' behavioral attitude, subjective norm, and perceived behavioral control (H1a–H1c are supported). The results suggest that AI literacy is not merely a technical skill but also a cognitive and psychological resource. Higher AI literacy helps teachers deepen their understanding of the educational value of AI, fostering the formation of a positive attitude (Ivanov et al., 2024; Ji et al., 2025). As a multidimensional construct, AI Literacy shapes teachers' psychological and cognitive readiness to embrace new technologies. By understanding AI's potential in education, teachers are more likely to develop positive attitudes toward its use, which, in turn, enhances their willingness to engage in innovative teaching practices. The strengthening of subjective norm indicates that teachers with higher literacy are more likely to recognize and internalize the expectations from external sources, such as policies, colleagues, and students, which further reinforces their professional responsibilities (Dierendonck et al., 2024). Aligns with TPB (Ajzen, 1991), these findings highlight the role of external expectations in shaping behavior. Teachers with higher AI literacy are not only more attuned to these external norms but are also more likely to integrate them into their professional identity, driving their engagement with AI in teaching. This emphasizes the significant role of external pressures and institutional support in facilitating teaching innovation. The enhancement of perceived behavioral control shows that AI literacy boosts self-efficacy, enabling teachers to navigate challenges and use AI tools in their teaching practices, thereby creating a mechanism of "technological mastery—efficacy improvement—behavioral transformation" (Viberg et al., 2024). Such findings underline the critical role of self-efficacy in fostering teaching innovation. When teachers feel competent and confident in using AI tools, they are more likely to engage in experimental and innovative behaviors, breaking free from traditional teaching models. This supports Wang and Zhao (2021), who emphasize the central role of self-efficacy in translating knowledge and skills into actual behaviors.

Behavioral attitude, subjective norm, and perceived behavioral control all significantly and positively predict teaching innovation (H2a–H2c are supported). This finding further validates the importance of these three psychological factors in the Theory of Planned Behavior (TPB) for translating intentions into actual behaviors. Perceived behavioral control had the most significant effect, indicating that when teachers feel confident in mastering and using AI, they are more likely to move away from traditional models and experiment with new practices. Specifically, when teachers feel equipped with the necessary skills and confidence to handle challenges, they are more likely to break free from conventional methods and engage in innovative behaviors. This aligns with the findings of Opoku et al. (2021) and Ramnarain et al. (2024), confirming the central role of self-efficacy in teaching innovation. A positive behavioral attitude reflects the recognition of AI's educational value, which in turn translates into motivation for innovation. Teachers, who understand the potential of AI in education, are more inclined to incorporate AI into their teaching practices. Granström and Oppi (2025) emphasize that teachers' positive attitudes are not just about technical proficiency but also about the recognition of AI's broader educational value, which drives them to innovate. The formation of such an attitude is underpinned by a shift from mere technical acceptance to a deeper understanding of the educational benefits, providing teachers with the motivation needed to embrace innovation. The impact of subjective norms shows that when teachers feel external expectations, they perceive innovation as an essential way to fulfill their professional roles and responsibilities. Policy support, peer expectations, and student demands play key roles in driving teachers' engagement with innovation (Cai and Tang, 2022). This finding highlights the significant influence of external pressures on shaping teachers' behavior. Teachers not only internalize these external expectations but also integrate them into their professional identity, reinforcing their commitment to adopting AI in their teaching practices. This underscores the role of institutional support and societal norms in facilitating teaching innovation.

Mediation analysis reveals that behavioral attitude, subjective norm, and perceived behavioral control all partially mediate the relationship between AI literacy and teaching innovation (H6a–H6c are supported). Among these, perceived behavioral control emerged as the most significant mediator, emphasizing the central role of self-efficacy in translating AI literacy into practical teaching innovation. This finding underscores the idea that teachers who feel confident in their ability to use AI tools are more likely to engage in innovative behaviors. As noted by Wang and Zhao (2021), self-efficacy plays a critical role in bridging the gap between knowledge acquisition and actual behavioral change, making it a key factor for fostering teaching innovation. While behavioral attitude and subjective norm also mediate the relationship, their effects were comparatively weaker. Behavioral attitude, which reflects teachers' recognition of AI's educational value, plays an important role in motivating innovation. However, without sufficient confidence in AI's practical application, sustaining innovation becomes challenging. As Peng et al. (2024) argue, positive attitudes alone are not enough to overcome barriers to adoption. Teachers must feel equipped with the necessary skills and support to translate their recognition of AI's value into consistent and meaningful teaching practices. The subjective norm, which relates to external expectations from peers, policies, and institutional pressures, can also play a role in promoting innovation. However, its mediating effect is more limited. Relying solely on external pressure can lead to compliance-based innovation, where teachers adopt new methods only because they feel obligated rather than motivated by a genuine desire for exploration and growth (Lu and Wang, 2023). Such innovations are more likely to be superficial and short-lived, as they lack intrinsic motivation or autonomy.

TABLE 7 Explanatory power (R^2) and predictive relevance (Q^2) of the structural model.

Endogenous latent variables	R^2	Q^2
BA	0.284	0.205
SN	0.269	0.190
PBC	0.309	0.224
TI	0.299	0.190

Source(s): created by author.

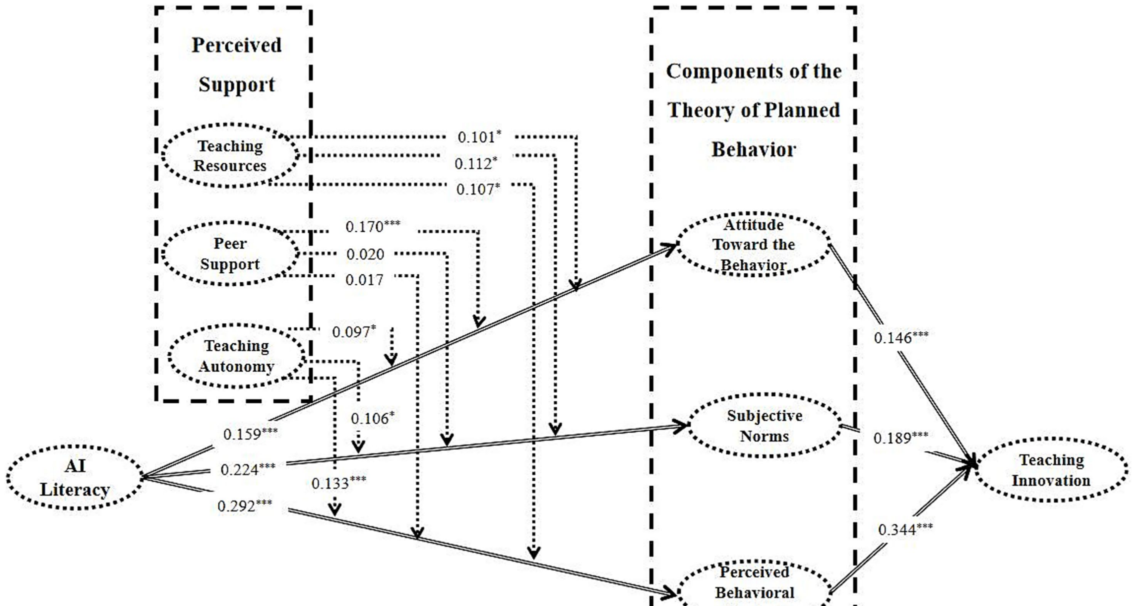
Moderating effect analysis shows that perceived support conditions play an important role in the "AI literacy—psychological mechanism" path. Teaching resources exhibited significant positive moderating effects on all three paths (H3a–H3c are supported). The availability of resources provides teachers with necessary tools, technical training, and institutional support, making it easier for AI literacy to translate into positive attitudes, norms, and control beliefs (Hazzan-Bishara et al., 2025). The effect of peer support was selective, being significant only in the relationship between AI literacy and behavioral attitude (H4a is supported), with no significant effects on subjective norm and perceived behavioral control (H4b and H4c are not supported). Attitude formation relies on emotional resonance and value recognition, with peers providing psychological support through experience sharing and belief dissemination (Habibi et al., 2023). Subjective norms are more shaped by policy guidance and institutional requirements, and informal peer opinions are less likely to serve as primary reference points (Dierendonck et al., 2024). Perceived behavioral control relies on teachers' self-confirmation of their abilities, with efficacy developed through accumulated experience, technical mastery, and teaching feedback, thus limiting the role of peer support (Gordon et al., 2023). University teachers typically have high professional autonomy, relying more on institutional signals and personal experience than on peer opinions for normative cognition and ability judgment. Therefore, the influence of peer support is concentrated in the attitude dimension, with boundaries in the formation of normative cognition and efficacy. Teaching autonomy exhibited significant positive moderating effects on all three paths (H5a–H5c are supported). In high-autonomy environments, teachers have greater decision-making power and freedom to experiment, allowing them to flexibly integrate AI tools in teaching design and practice. Self-determination theory (Deci and Ryan, 2000) suggests that autonomy can stimulate intrinsic motivation and exploratory desire, while Hou and Shen (2024) further emphasize its role in promoting responsibility and sustainability. The study results confirm the critical role of teaching autonomy in transforming AI literacy into innovative behavior.

7 Contributions, limitations, and future research directions

7.1 Theoretical contributions

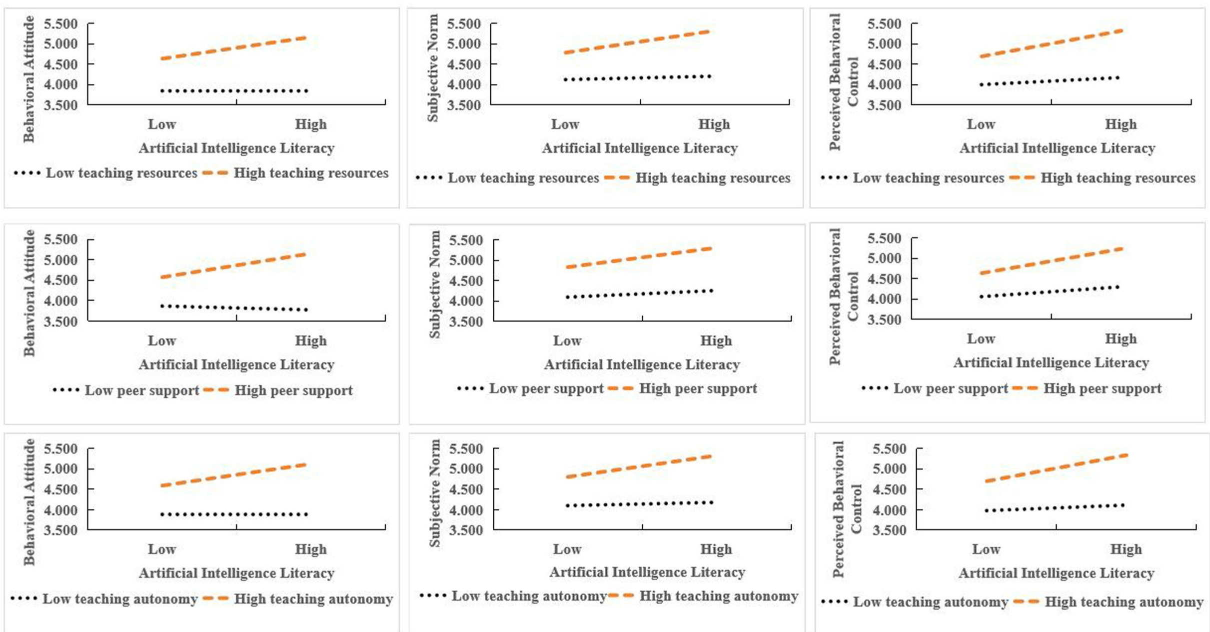
This study develops a framework based on the Theory of Planned Behavior (TPB) to explore how AI literacy influences teaching innovation through psychological mechanisms such as behavioral attitude, subjective norm, and perceived behavioral control. The research also introduces perceived support factors such as teaching resources, peer support, and teaching autonomy, further revealing their moderating roles between AI literacy and teaching innovation. The theoretical contributions of this study are as follows:

First, it expands the psychological connotations of AI literacy. Existing studies often regard AI literacy as an external manifestation of technical abilities and tool usage (Ng et al., 2021), while this research redefines AI literacy from a psychological perspective. The study finds that AI literacy is not just a technical competence but also a cognitive and psychological resource that influences teachers'



Source(s): Created by author

FIGURE 3
Structural model results.



Source(s): Created by author

FIGURE 4
Simple slope plot.

behavioral attitude, subjective norm, and perceived behavioral control. Teachers with higher AI literacy are more likely to recognize the potential of AI in education, forming a positive teaching attitude. This finding provides a new definition for the theoretical system of AI literacy and reveals its multidimensional impact on teachers' psychology and behavior (Ayanwale et al., 2024; Sperling et al., 2024).

Second, it deepens the application of TPB in educational technology adoption. This study enriches the application of TPB in the educational field by verifying the roles of behavioral attitude, subjective norm, and perceived behavioral control in teaching innovation. It reveals how AI literacy influences teachers' innovative behaviors through these three dimensions. Specifically, the study

TABLE 8 Moderating effect at different levels.

Independent variable	Dependent variable	Moderator variable	TA	β	95%CI		p-value
AIL	BA	TR	Low	−0.008	−0.165	0.149	0.923
			Medium	0.262	0.165	0.359	0.000***
			High	0.531	0.412	0.650	0.000***
	SN		Low	0.083	−0.071	0.236	0.291
			Medium	0.310	0.215	0.405	0.000***
			High	0.538	0.421	0.654	0.000***
	PBC		Low	0.174	0.021	0.328	0.026*
			Medium	0.401	0.306	0.496	0.000***
			High	0.628	0.511	0.744	0.000***
	BA	PS	Low	−0.093	−0.260	0.073	0.270
			Medium	0.245	0.147	0.343	0.000***
			High	0.584	0.460	0.708	0.000***
	SN		Low	0.162	−0.001	0.325	0.052
			Medium	0.320	0.224	0.416	0.000***
			High	0.478	0.356	0.600	0.000***
	PBC		Low	0.241	0.076	0.407	0.004**
			Medium	0.418	0.320	0.516	0.000***
			High	0.595	0.471	0.719	0.000***
	BA	TA	Low	−0.014	−0.179	0.150	0.863
			Medium	0.259	0.159	0.359	0.000***
			High	0.533	0.413	0.652	0.000***
	SN		Low	0.081	−0.078	0.240	0.318
			Medium	0.304	0.207	0.400	0.000***
			High	0.527	0.411	0.642	0.000***
	PBC		Low	0.135	−0.023	0.293	0.094
			Medium	0.387	0.291	0.483	0.000***
			High	0.639	0.524	0.753	0.000***

*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.
Source(s): created by author.

TABLE 9 Results of mediating effect analysis.

Mediating path	Effect value	S. E	Lower 2.5%	Upper 2.5%	p
AIL→BA→TI	0.023	0.010	0.007	0.045	***
AIL→SN → TI	0.042	0.013	0.020	0.070	***
AIL→PBC → TI	0.101	0.018	0.066	0.138	***
AIL→TI	0.166	0.022	0.125	0.211	***

***$p < 0.001$.
Source(s): created by author.

demonstrates that perceived behavioral control (self-efficacy) plays a central role in teachers' innovative behaviors. AI literacy enhances self-efficacy, helping teachers overcome technological challenges and promote teaching innovation (Frawley and Campbell, 2025; Gold et al., 2024). This finding not only deepens the theoretical application of TPB in educational technology adoption but also provides a new theoretical framework for understanding teachers' innovative behaviors (Bali et al., 2025).

Third, it fills the gap in existing theories by incorporating the Theory of Planned Behavior into the study of teaching innovation. Previous theories on technology adoption and teaching innovation often overlook the psychological mechanisms that drive teachers' willingness to adopt new tools and engage in innovative behaviors. For instance, traditional models have often failed to explicitly account for the role of subjective norms and perceived behavioral control in influencing teachers' innovative actions. By integrating these components of TPB, this study provides a more comprehensive understanding of how teachers' psychological states—shaped by both internal capabilities (e.g., attitudes) and external expectations (e.g., norms)—interact to influence their engagement in innovation (Ma and Lei, 2024). TPB's inclusion of both cognitive and social dimensions of decision-making offers a more robust theoretical framework for analyzing educational innovation.

Finally, it explores the moderating role of perceived support factors in teaching innovation. By incorporating teaching resources, peer support, and teaching autonomy into the TPB framework, this study investigates their moderating effects between AI literacy and teaching innovation. The results show that adequate teaching resources, effective peer support, and higher teaching autonomy significantly enhance the impact of AI literacy on teachers' psychological mechanisms, further promoting teaching innovation (Ramnarain et al., 2024). This finding provides new insights into the internal and external drivers of teachers' innovative behaviors, expands the application boundaries of TPB, and offers theoretical support for future educational policies and teacher training designs (Cao et al., 2022).

7.2 Practical implications

This study investigates the mechanisms through which university teachers' AI literacy influences teaching innovation, providing several implications for educational practice and policy-making.

First, teacher training should go beyond basic technical operation and redefine AI literacy as a comprehensive capability encompassing both cognitive and psychological aspects. Universities can design modular courses and practical seminars to help teachers master the application of tools, while reinforcing educational value recognition and critical thinking through case analysis and scenario exercises, thus enhancing overall literacy on both skills and psychological levels.

Second, teaching innovation depends on the enhancement of teachers' perceived behavioral control. Administrators should foster self-efficacy through continuous feedback, progressive tasks, and simulated teaching, allowing teachers to maintain confidence in uncertain and challenging situations. Higher levels of perceived control can translate into stable innovative intentions and practices.

Third, perceived support is a crucial condition for fostering teachers' innovation. Schools must ensure the availability of educational resources, including technical training, digital platforms, and interdisciplinary collaboration opportunities, to strengthen teaching preparation. Peer support can be implemented through academic community building, experience sharing, and collaborative projects, providing emotional support on the value and affective levels. Teaching autonomy should be guaranteed through institutional arrangements, empowering teachers with decision-making authority in course design, tool selection, and teaching methods, thereby stimulating exploration motivation and continuous innovation.

Finally, education policymakers should consider multilevel needs. National and regional policies should incentivize the balanced distribution of AI education resources; at the school level, layered training should be designed based on teachers' professional development stages; at the individual level, flexible autonomy and continuous support should be provided to guide teachers in transforming AI literacy into visible teaching innovation practices.

7.3 Research limitations

The data in this study were sourced from university teachers in China, and the sample is concentrated in terms of geographic and institutional backgrounds, limiting the generalizability of the findings across different cultures and institutional contexts. The research design used a cross-sectional survey, which makes it difficult to fully validate causal relationships between variables. The research tool mainly relied on self-reported questionnaires, which may have introduced social desirability bias and subjective bias. The dimensions of perceived support factors were relatively limited, focusing only on teaching resources, peer support, and teaching autonomy, without addressing broader contextual variables such as leadership support, organizational climate, and educational policies. The model also lacked a thorough examination of differences among teacher groups, with insufficient exploration of heterogeneity across disciplines and career stages. The research methodology predominantly used quantitative analysis, leaving limited space for capturing teachers' real psychological experiences and practical logics.

7.4 Future research directions

Future research could expand the sample to include university teachers from different countries and regions to test the universality and differences of AI literacy across diverse cultural and institutional contexts. Longitudinal tracking and experimental designs could be employed to observe the development of teachers' literacy and innovation pathways over time. The research dimensions of perceived support should be extended to include leadership support, organizational climate, and educational policies, constructing a more comprehensive contextual framework. Future studies could also focus on the heterogeneity of teacher groups, comparing differences in the mechanisms across disciplines and career stages, and revealing the interactive effects between group and individual factors. The research methodology could combine both qualitative and quantitative approaches, using interviews, classroom observations, and case studies to gain a deeper understanding of teachers' psychological experiences and practical logics in teaching innovation.

Data availability statement

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

Author contributions

YZ: Project administration, Funding acquisition, Writing – review & editing, Resources, Validation, Supervision, Formal analysis, Software, Writing – original draft, Data curation, Investigation, Methodology, Visualization, Conceptualization. LH: Software, Project administration, Funding acquisition, Writing – review & editing, Resources, Formal analysis, Writing – original draft, Supervision, Data curation, Methodology, Visualization, Conceptualization, Investigation, Validation.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This work was supported by the Research on the Construction of the Evaluation Indicator System for Online Courses of Domestic College Students: A Key

Project of the “13th Five-Year Plan” for Educational Science in Beijing (2020) (Grant No. AHAA2020052).

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.

Generative AI statement

The authors declare that no Gen AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

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

Abulibdeh, A., Zaidan, E. S., and Abulibdeh, R. (2024). Navigating the confluence of artificial intelligence and education for sustainable development in the era of industry 4.0: challenges, opportunities, and ethical dimensions. *J. Clean. Prod.* 437:140527. doi: 10.1016/j.jclepro.2023.140527

Adabor, E. S., Addy, E., Assyne, N., and Antwi-Boasiako, E. (2025). Enhancing sustainable academic course delivery using AI in technical universities: an empirical analysis using adaptive learning theory. *Sustain. Futures* 10:100828. doi: 10.1016/j.sftr.2025.100828

Adams, D., Bellibas, M. S., Muniandy, V., and Pietsch, M. (2025). The role of openness to experience in innovating teaching and instruction through leader-member exchange and teacher creativity. *Think. Skills Creat.* 57:101834. doi: 10.1016/j.tsc.2025.101834

Adelana, O. P., Ayanwale, M. A., and Sanusi, I. T. (2024). Exploring pre-service biology teachers’ intention to teach genetics using an AI intelligent tutoring - based system. *Cogent Educ.* 11:2310976. doi: 10.1080/2331186X.2024.2310976

Adie, J. W., Waldeck, D., Hughes, D. J., Mulvenna, M., Holliman, A. J., Bartholomew, K. J., et al. (2024). Peer climate matters for academic motivation and student functioning in higher education. *Int. J. Appl. Posit. Psychol.* 9, 1685–1702. doi: 10.1007/s41042-024-00187-w

Ahadzadeh, A. S., Wu, S. L., and Sijia, X. (2024). Exploring academic intentions for ChatGPT: a perspective from the theory of planned behavior. *CMU J. Soc. Sci. Human.* 11:e2024016. doi: 10.12982/CMUJASR.2024.016

Ajzen, I. (1991). The theory of planned behavior. *Organ. Behav. Hum. Decis. Process.* 50, 179–211. doi: 10.1016/0749-5978(91)90020-T

Almatrafi, O., Johri, A., and Lee, H. (2024). A systematic review of AI literacy conceptualization, constructs, and implementation and assessment efforts (2019–2023). *Comput. Educ. Open.* 6:100173. doi: 10.1016/j.caeo.2024.100173

Al-Mughairi, H., and Bhaskar, P. (2024). Exploring the factors affecting the adoption AI techniques in higher education: insights from teachers' perspectives on ChatGPT. *J. Res. Innovative Teach. Learn.* 18, 232–247. doi: 10.1108/JRIT-09-2023-0129

Andersen, H. K., Mayerl, J., Hornung, G., and Thyssen, C. (2019). Attitude-behaviour relations in teaching natural science. Results from a longitudinal study using the theory of planned behaviour. *SocArXiv*. 1–40. doi: 10.31235/osf.io/dntw2

Ateş, H. (2020). Merging theory of planned behavior and value identity personal norm model to explain pro-environmental behaviors. *Sustain. Prod. Consum.* 24, 169–180. doi: 10.1016/j.spc.2020.07.006

Ayanwale, M. A., Adelana, O. P., Bamiro, N. B., Olatunbosun, S. O., Idowu, K. O., and Adewale, K. A. (2025). Large language models and GenAI in education: insights from Nigerian in-service teachers through a hybrid ANN-PLS-SEM approach. *F1000Res* 14:258. doi: 10.12688/f1000research.161637.1

Ayanwale, M. A., Adelana, O. P., Molefi, R. R., Adeeko, O., and Ishola, A. M. (2024). Examining artificial intelligence literacy among pre-service teachers for future classrooms. *Comp. Educ. Open* 6:100179. doi: 10.1016/j.caeo.2024.100179

Bali, S., Chen, T.-C., and Liu, M.-C. (2025). Behavioral intentions of low-achieving students to use mobile english learning: integrating self-determination theory, theory of planned behavior, and technology acceptance model approaches. *Int. J. Human–Computer Interaction* 41, 5522–5532. doi: 10.1080/10447318.2024.2364142

Bao, Y. G. (2024). Principals' secure base leadership and teachers' innovative behavior: a moderated mediation model. *Soc. Behav. Pers.* 52:e13121. doi: 10.2224/sbp.13121

Bewersdorff, A., Hornberger, M., Nerdel, C., and Schiff, D. S. (2025). AI advocates and cautious critics: how AI attitudes, AI interest, use of AI, and AI literacy build university students' AI self-efficacy. *Comp. Educ. Artificial Intell* 8:100340. doi: 10.1016/j.caeai.2024.100340

Broadbent, J., Bearman, M., Boud, D., and Dawson, P. (2024). Academics’ intention to sustain new teaching practices after the COVID-19 pandemic: examined through the theory of planned behaviour. *High. Educ.* 90, 371–386. doi: 10.1007/s10734-024-01326-7

Cai, Y. H., and Tang, R. J. (2021). School support for teacher innovation: mediating effects of teacher self-efficacy and moderating effects of trust. *Think. Skills Creat.* 41:100854. doi: 10.1016/j.tsc.2021.100854

Cai, Y., and Tang, R. (2022). School support for teacher innovation: the role of basic psychological need satisfaction. *Think. Skills Creat.* 45:101096. doi: 10.1016/j.tsc.2022.101096

Cao, C., Chen, B., Yang, S., Zheng, X., Ye, Y., and Yue, X. (2022). Exploring the impact of individual and social antecedents on teachers’ teaching innovation: perspective of goal-oriented behavior and social identity [original research]. *Front. Psychol.* 13:924052. doi: 10.3389/fpsyg.2022.924052

Carrión, G. C., Nitzl, C., and Roldán, J. L. (2017). “Mediation analyses in partial least squares structural equation modeling: guidelines and empirical examples” in Partial least squares path Modeling: Basic concepts, methodological issues and applications. eds. H. Latan and R. Noonan (Cham: Springer International Publishing), 173–195.

Celik, I., Muukkonen, H., and Siklander, S. (2025). Teacher–artificial intelligence (AI) interaction: the role of trust, subjective norm and innovativeness in Teachersacceptance of educational chatbots. *Policy Futures Educ.* 14782103251348551. doi: 10.1177/1478210325134

Chen, M.-H., Agrawal, S., Lin, S.-M., and Liang, W.-L. (2023). Learning communities, social media, and learning performance: transactive memory system perspective. *Comput. Educ.* 203:104845. doi: 10.1016/j.compedu.2023.104845

Chen, J. H., He, X., Ning, Y. M., Wijaya, T. T., and Liu, J. H. (2025). Task/technology fit or technology attraction? The intentions of STEM teachers to use AI technologies for teaching innovation. *Asia-Pacific Educ. Res.* 34, 1–15. doi: 10.1007/s40299-025-01020-x

Chen, Y. (2024). Teacher autonomy, creative self-efficacy, and innovative behavior: perspectives from Chinese university EFL teachers. Arab world. *Engl. J.* 15, 73–86. doi: 10.24093/awej/vol15no2.5

Chen, Y. Y., and Zou, Y. T. (2024). Enhancing education quality: exploring teachers' attitudes and intentions towards intelligent MR devices. *Eur. J. Educ.* 59:12692. doi: 10.1111/ejed.12692

Cheung, G. W., Cooper-Thomas, H. D., Lau, R. S., and Wang, L. C. (2024). Reporting reliability, convergent and discriminant validity with structural equation modeling: a review and best-practice recommendations. *Asia Pac. J. Manag.* 41, 745–783. doi: 10.1007/s10490-023-09871-y

Chou, C. M., Shen, T. C., Shen, T. C., and Shen, C. H. (2025). Teachers' adoption of AI-supported teaching behavior and its influencing factors: using structural equation modeling. *J. Comput. Educ.* 12, 853–896. doi: 10.1007/s40692-024-00332-z

Chu, T., and Wang, W. Y. (2025). Exploring EFL teachers' epistemic agency and its influencing factors in an AI-integrated context through the double stimulation model. *Asia-Pacific Education Res.* 34, 1267–1275. doi: 10.1007/s40299-024-00940-4

Cui, C., and Yin, Y. J. (2023). How professional training impacts teaching innovation among ideological and political teachers: the mediating and moderating role of basic psychological needs satisfaction. *Front. Psychol.* 14:1246951. doi: 10.3389/fpsyg.2023.1246951

Deci, E. L., and Ryan, R. M. (2000). The "what" and "why" of goal pursuits: human needs and the self-determination of behavior. *Psychol. Inq.* 11, 227–268. doi: 10.1207/s15327965pli1104_01

Dierendonck, C., Poncelet, D., and Tinnes-Vigne, M. (2024). Why teachers do (or do not) implement recommended teaching practices? An application of the theory of planned behavior. *Front. Psychol.* 15:1269954. doi: 10.3389/fpsyg.2024.1269954

Ding, A.-C. E., Shi, L., Yang, H., and Choi, I. (2024). Enhancing teacher AI literacy and integration through different types of cases in teacher professional development. *Comput. Educ. Open* 6:100178. doi: 10.1016/j.caeo.2024.100178

Dunn, R., Hattie, J., and Bowles, T. (2018). Using the theory of planned behavior to explore teachers' intentions to engage in ongoing teacher professional learning. *Stud. Educ. Eval.* 59, 288–294. doi: 10.1016/j.stueduc.2018.10.001

Duong, C. D. (2025). AI literacy and e-entrepreneurial intention: a serial mediation model of e-entrepreneurial self-efficacy and e-entrepreneurial identity aspiration. *Int. J. Inf. Manag. Data Insights* 5:100349. doi: 10.1016/j.jjimei.2025.100349

Ehlert, M., Hebbecker, K., and Souvignier, E. (2022). Implementing evidence-based practices in reading classrooms – longitudinal interrelations between teaching-related attitudes, innovation-specific perceptions, and perceived implementation success. *Stud. Educ. Eval.* 75:101203. doi: 10.1016/j.stueduc.2022.101203

Ertas, B. D., and Pekmezci, F. B. (2025). Career motivation and job satisfaction in Türkiye: mediating role of teacher innovativeness and instructional practice. *Asia-Pacific Educ. Res.* 34, 193–207. doi: 10.1007/s40299-024-00846-1

Frawley, C., and Campbell, L. O. (2025). Factors that predict teachers' intentions to utilize emerging technologies: an investigation using PLS-SEM. *Educ. Inf. Technol.* 30, 1589–1606. doi: 10.1007/s10639-024-12796-8

Gerçek, M., and Özveren, C. G. (2025). Creative self-efficacy, learning agility, and proactive career behaviors: the moderated mediation effect of positive AI attitudes among young adults. *Think. Skills Creat.* 58:101895. doi: 10.1016/j.tsc.2025.101895

Gilbert, A., Tait-McCutcheon, S., and Knewstubb, B. (2021). Innovative teaching in higher education: teachers' perceptions of support and constraint. *Innov. Educ. Teach. Int.* 58, 123–134. doi: 10.1080/14703297.2020.1715816

Gold, B., Thomm, E., and Bauer, J. (2024). Using the theory of planned behaviour to predict pre-service teachers' preferences for scientific sources. *Br. J. Educ. Psychol.* 94, 216–230. doi: 10.1111/bjep.12643

Gong, L. (2023). How can college teachers' information literacy empower their innovative teaching behaviors under the background of informatization? *PLoS One* 18:e0294593. doi: 10.1371/journal.pone.0294593

Gordon, D., Blundell, C., Mills, R., and Bourke, T. (2023). Teacher self-efficacy and reform: a systematic literature review. *Aust. Educ. Res.* 50, 801–821. doi: 10.1007/s13384-022-00526-3

Granström, M., and Oppi, P. (2025). Assessing teachers' readiness and perceived usefulness of AI in education: an Estonian perspective [original research]. *Front. Educ.* 10:1622240. doi: 10.3389/feduc.2025.1622240

Guan, L., Zhang, Y., and Gu, M. M. (2025). Pre-service teachers preparedness for AI-integrated education: an investigation from perceptions, capabilities, and teachers' identity changes. *Comp. Educ. Artificial Intell.* 8:100341. doi: 10.1016/j.caeai.2024.100341

Gupta, K. P., and Bhaskar, P. (2020). Inhibiting and motivating factors influencing teachers' adoption of AI-based teaching and learning solutions: prioritization using analytic hierarchy process. *J. Inf. Technol. Educ. Res.* 19:693. doi: 10.28945/4640

Habibi, A., Riady, Y., Samed Al-Adwan, A., and Awni Albelbisi, N. (2023). Beliefs and knowledge for pre-service teachers' technology integration during teaching practice: an extended theory of planned behavior. *Comput. Sch.* 40, 107–132. doi: 10.1080/07380569.2022.2124752

Hamm, J. M., Lachman, M. E., Duggan, K. A., Mogle, J. A., McGrath, R., Parker, K., et al. (2024). When and how perceived control buffers against cognitive declines: a moderated mediation analysis. *Psychol. Aging.* 40, 39–53. doi: 10.1037/pag0000841

Han, J., Gao, C., and Yang, J. (2021). Chinese university efl teachers' perceived support, innovation, and teaching satisfaction in online teaching environments: the mediation of teaching efficacy. *Front. Psychol.* 12:761106. doi: 10.3389/fpsyg.2021.761106

Hayes, A. F. (2009). Beyond baron and kenny: statistical mediation analysis in the new millennium. *Commun. Monogr.* 76, 408–420. doi: 10.1080/03637750903310360

Hazzan-Bishara, A., Kol, O., and Levy, S. (2025). The factors affecting teachers' adoption of AI technologies: a unified model of external and internal determinants. *Educ. Inf. Technol.* 30, 15043–15069. doi: 10.1007/s10639-025-13393-z

Henseler, J., Ringle, C. M., and Sarstedt, M. (2015). 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

Hornstra, L., Stroet, K., and Weijers, D. (2021). Profiles of teachers' need-support: how do autonomy support, structure, and involvement cohere and predict motivation and learning outcomes? *Teach. Teach. Educ.* 99:103257. doi: 10.1016/j.tate.2020.103257

Hou, M., Lin, Y., Shen, Y., and Zhou, H. (2022). Explaining pre-service teachers' intentions to use technology-enabled learning: an extended model of the theory of planned behavior. *Front. Psychol.* 13:900806. doi: 10.3389/fpsyg.2022.900806

Hou, M., and Shen, Y. (2024). Explaining preservice teachers' intention and behavior to use technology-enabled learning in China: a multi-group analysis across experiences. *Psychol. Schs.* 61, 4538–4557. doi: 10.1002/pits.23295

Hwang, Y., and Wu, Y. (2025). The influence of generative artificial intelligence on creative cognition of design students: a chain mediation model of self-efficacy and anxiety [original research]. *Front. Psychol.* 15:1455015. doi: 10.3389/fpsyg.2024.1455015

Hwang, G.-J., Xie, H., Wah, B. W., and Gašević, D. (2020). Vision, challenges, roles and research issues of artificial intelligence in education. *Comp. Educ. Artificial Intell.* 1:100001. doi: 10.1016/j.caeai.2020.100001

Ivanov, S., Soliman, M., Tuomi, A., Alkathiri, N. A., and Al-Alawi, A. N. (2024). Drivers of generative AI adoption in higher education through the lens of the theory of planned behaviour. *Technol. Soc.* 77:102521. doi: 10.1016/j.techsoc.2024.102521

Ji, Y., Zhong, M., Lyu, S., Li, T., Niu, S., and Zhan, Z. (2025). How does AI literacy affect individual innovative behavior: the mediating role of psychological need satisfaction, creative self-efficacy, and self-regulated learning. *Educ. Inf. Technol.* 30, 16133–16162. doi: 10.1007/s10639-025-13437-4

Jose, B., Cherian, J., Verghis, A. M., Varghise, S. M., Mumthas, S., and Joseph, S. (2025). The cognitive paradox of AI in education: between enhancement and erosion [opinion]. *Front. Psychol.* 16:1550621. doi: 10.3389/fpsyg.2025.1550621

Kelley, M., and Wenzel, T. (2025). Advancing artificial intelligence literacy in teacher education through professional partnership inquiry. *Educ. Sci.* 15:659. doi: 10.3390/educsci15060659

Kim, J. (2024). Leading teachers' perspective on teacher-AI collaboration in education. *Educ. Inf. Technol.* 29, 8693–8724. doi: 10.1007/s10639-023-12109-5

Kim, A. (2025). Exploring how institutional structures shape instructional quality: the roles of teacher education, employment, and development in fostering higher-order thinking skills. *Think. Skills Creat.* 58:101954. doi: 10.1016/j.tsc.2025.101954

Knauder, H., and Koschmieder, C. (2019). Individualized student support in primary school teaching: a review of influencing factors using the theory of planned behavior (TPB). *Teach. Teach. Educ.* 77, 66–76. doi: 10.1016/j.tate.2018.09.012

Kock, N. (2015). Common method bias in PLS-SEM: a full collinearity assessment approach. *Int. J. e-Collaboration (ijec)* 11, 1–10. doi: 10.4018/ijec.2015100101

Kong, S. C., Yang, Y., and Hou, C. (2024). Examining teachers' behavioural intention of using generative artificial intelligence tools for teaching and learning based on the extended technology acceptance model. *Comp. Educ. Artificial Intell.* 7:100328. doi: 10.1016/j.caeai.2024.100328

Kottmann, A., Schildkamp, K., and van der Meulen, B. (2024). Determinants of the innovation behaviour of teachers in higher education. *Innov. High. Educ.* 49, 397–418. doi: 10.1007/s10755-023-09689-y

Kruse, F., Büchel, S., and Brühwiler, C. (2024). Longitudinal effects of basic psychological need support on the development of intrinsic motivation and perceived competence in physical education. A multilevel study. *Front. Psychol.* 15:1393966. doi: 10.3389/fpsyg.2024.1393966

Liao, K., Liu, Z. Y., and Li, B. (2022). The effect of psychological capital and role conflict on the academic entrepreneurial intents of Chinese teachers in higher education: a study based on the theory of planned behavior. *Front. Psychol.* 13:793408. doi: 10.3389/fpsyg.2022.793408

Lim, E. M. (2023). The effects of pre-service early childhood teachers' digital literacy and self-efficacy on their perception of AI education for young children. *Educ. Inf. Technol.* 28, 12969–12995. doi: 10.1007/s10639-023-11724-6

Lin, H. Z., and Chen, Q. (2024). Artificial intelligence (AI) -integrated educational applications and college students' creativity and academic emotions: students and teachers' perceptions and attitudes. *BMC Psychol.* 12:487. doi: 10.1186/s40359-024-01979-0

Liu, J., and Chang, Y.-C. (2023). The influence of physical education teachers' perceived organization support on innovation behavior: the mediating role of learning goal orientation. *Educ. Res. Rev.* 18, 260–268. doi: 10.5897/ERR2023.4349

Liu, H., Fan, J., and Xia, M. (2025). Exploring individual's emotional and autonomous learning profiles in AI-enhanced data-driven language learning: an expanded sor perspective. *Learn. Individ. Differ.* 122:102753. doi: 10.1016/j.lindif.2025.102753

Liu, S., Hallinger, P., and Feng, D. (2016). Supporting the professional learning of teachers in China: does principal leadership make a difference? *Teach. Teach. Educ.* 59, 79–91. doi: 10.1016/j.tate.2016.05.023

Liu, S. N., Lu, J. F., and Yin, H. B. (2022). Can professional learning communities promote teacher innovation? A multilevel moderated mediation analysis. *Teach. Teach. Educ.* 109:103571. doi: 10.1016/j.tate.2021.103571

Liu, G. L., and Wang, Y. (2024). Modeling EFL teachers' intention to integrate informal digital learning of English (IDLE) into the classroom using the theory of planned behavior. *System* 120:103193. doi: 10.1016/j.system.2023.103193

Liu, S. N., Yin, H. B., Wang, Y. N., and Lu, J. F. (2024). Teacher innovation: conceptualizations, methodologies, and theoretical framework. *Teach. Teach. Educ.* 145:104611. doi: 10.1016/j.tate.2024.104611

López, C. M., Saborido, D. M., Seijo, J. C. T., and Batanero, J. M. F. (2023). Explanatory variables of educational innovation: a model based on personality traits. *Electron. J. Res. Educ. Psychol.* 21, 349–372. doi: 10.25115/ejrep.v21i60.5538

Lu, H. P., and Wang, J. C. (2023). Exploring the effects of sudden institutional coercive pressure on digital transformation in colleges from teachers' perspective. *Educ. Inf. Technol.* 28, 15991–16015. doi: 10.1007/s10639-023-11781-x

Ma, Y. (2025). Modelling college students' acceptance to use generative artificial intelligence for second language learning: a theory of planned behaviour perspective. *Eur. J. Educ.* 60:e12923. doi: 10.1111/ejed.12923

Ma, S., and Lei, L. (2024). The factors influencing teacher education students' willingness to adopt artificial intelligence technology for information-based teaching. *Asia Pac. J. Educ.* 44, 94–111. doi: 10.1080/02188791.2024.2305155

Ma, X. W., and Zhang, P. (2025). How distributed leadership shapes chinese teachers' teaching innovation behavior: mediating roles of teacher collaboration and teachers' social-emotional learning beliefs. *Asia-Pacific Educ. Res.* 34, 1–13. doi: 10.1007/s40299-025-01006-9

Martinek, D., Zumbach, J., and Carmignola, M. (2020). The impact of perceived autonomy support and autonomy orientation on orientations towards teaching and self-regulation at university. *Int. J. Educ. Res.* 102:101574. doi: 10.1016/j.ijer.2020.101574

Miao, J., and Ma, L. (2023). Teacher autonomy support influence on online learning engagement: the mediating roles of self-efficacy and self-regulated learning. *SAGE Open* 13:21582440231217737. doi: 10.1177/21582440231217737

Miranda, J., Navarrete, C., Noguez, J., Molina-Espinosa, J.-M., Ramírez-Montoya, M.-S., Navarro-Tuch, S. A., et al. (2021). The core components of education 4.0 in higher education: Three case studies in engineering education. *Computers & Electrical Engineering* 93:107278. doi: 10.1016/j.compeleceng.2021.107278

Mnguni, L., Nuangchalerm, P., Zaky El Islami, R. A., Sibanda, D., Sari, I. J., and Ramulumo, M. (2024). The behavioural intentions for integrating artificial intelligence in science teaching among pre-service science teachers in South Africa and Thailand. *Comp. Educ. Artificial Intell.* 7:100334. doi: 10.1016/j.caeai.2024.100334

Mokhlis, S., and Abdullah, A. (2024). Teachers' autonomy and innovative work behavior: the mediating role of schools' innovation climate. *Int. J. Eval. Res. Educ.* 13, 2108–2115. doi: 10.11591/ijere.v13i4.27697

Ng, D. T. K., Leung, J. K. L., Chu, S. K. W., and Qiao, M. S. (2021). Conceptualizing AI literacy: an exploratory review. *Comp. Educ. Artificial Intell.* 2:100041. doi: 10.1016/j.caeai.2021.100041

Niloy, A. C., Akter, S., Sultana, J., Rahman, M. A., Sultana, N., Prome, T. I., et al. (2025). Can generative AI be an effective co-teacher? An experiment. *Comp. Educ. Artificial Intell.* 8:100418. doi: 10.1016/j.caeai.2025.100418

Ning, Y. M., Zheng, H. Y., Wu, H. D., Jin, Z. J., Chang, H. B., and Wijaya, T. T. (2025). Analysis of influencing factors on teachers' AI literacy under the SOR framework: an empirical study based on PLS-SEM and fsQCA. *Educ. Inf. Technol.* 30, 18213–18239. doi: 10.1007/s10639-025-13477-w

Nshimiyimana, A., and Cartledge, P. T. (2020). Peer-teaching at the University of Rwanda - a qualitative study based on self-determination theory. *BMC Med. Educ.* 20:230. doi: 10.1186/s12909-020-02142-0

Okada, R. (2023). Effects of perceived autonomy support on academic achievement and motivation among higher education students: a meta-analysis. *Jpn. Psychol. Res.* 65, 230–242. doi: 10.1111/jpr.12380

Opoku, M. P., Cuskelly, M., Pedersen, S. J., and Rayner, C. S. (2021). Applying the theory of planned behaviour in assessments of teachers' intentions towards practicing inclusive education: a scoping review. *Eur. J. Spec. Needs Educ.* 36, 577–592. doi: 10.1080/08856257.2020.1779979

Ozudogru, G., and Durak, H. Y. (2025). Conceptualizing pre-service teachers' artificial intelligence readiness and examining its relationship with various variables: the role of artificial intelligence literacy, digital citizenship, artificial intelligence-enhanced innovation and perceived threats from artificial intelligence. *Inf. Dev.* 41, 916–932. doi: 10.1177/02666669251335657

Padilha, J. M., Machado, P. P., Ribeiro, A. L., Ribeiro, R., Vieira, F., and Costa, P. (2021). Easiness, usefulness and intention to use a MOOC in nursing. *Nurse Educ. Today* 97:104705. doi: 10.1016/j.nedt.2020.104705

Panday-Shukla, P. (2025). Exploring generative artificial intelligence in teacher education. *Teach. Teach. Educ.* 165:105088. doi: 10.1016/j.tate.2025.105088

Peng, R., Razak, R. A., and Halili, S. H. (2024). Exploring the role of attitudes, self-efficacy, and digital competence in influencing teachers' integration of ICT: a partial least squares structural equation modeling study. *HELIYON* 10:e34234. doi: 10.1016/j.heliyon.2024.e34234

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, 879–903. doi: 10.1037/0021-9010.88.5.879

Qin, S. B., Jia, S. B., and Su, S. W. (2025). How does teacher collaboration impact teachers' innovation ability? The chain mediation of teaching motivation and teaching efficacy. Humanities & Social Sciences. *Communications* 12:924. doi: 10.1057/s41599-025-04965-y

Ramnarain, U., Ogegbo, A. A., Penn, M., Ojetunde, S., and Mdlalose, N. (2024). Pre-service science teachers' intention to use generative artificial intelligence in inquiry-based teaching. *J. Sci. Educ. Technol.* 33, 1–14. doi: 10.1007/s10956-024-10159-z

Reeve, J., and Cheon, S. H. (2024). Learning how to become an autonomy-supportive teacher begins with perspective taking: a randomized control trial and model test. *Teach. Teach. Educ.* 148:104702. doi: 10.1016/j.tate.2024.104702

Ren, X. (2024). Modeling college EFL teachers' intentions to conduct academic research: integrating theory of planned behavior with self-determination theory. *PLoS One* 19:e0307704. doi: 10.1371/journal.pone.0307704

Richter, S. L., and Schuessler, J. B. (2019). Nursing faculty experiences and perceptions of online teaching: a descriptive summary. *Teach. Learn. Nurs.* 14, 26–29. doi: 10.1016/j.teln.2018.09.004

Rigdon, E. E. (2016). Choosing PLS path modeling as analytical method in European management research: a realist perspective. *Eur. Manag. J.* 34, 598–605. doi: 10.1016/j.emj.2016.05.006

Robayo-Pinzon, O., Rojas-Berrio, S., Rincon-Novoa, J., and Ramirez-Barrera, A. (2024). Artificial intelligence and the value co-creation process in higher education institutions. *Int. J. Human-Computer Interaction* 40, 6659–6675. doi: 10.1080/10447318.2023.2259722

Sanusi, I. T., Ayanwale, M. A., and Tolorunleke, A. E. (2024). Investigating pre-service teachers' artificial intelligence perception from the perspective of planned behavior theory. *Comp. Educ. Artificial Intell.* 6:100202. doi: 10.1016/j.caeai.2024.100202

Sarstedt, M., Hair, J. F. Jr., Nitzl, C., Ringle, C. M., and Howard, M. C. (2020). Beyond a tandem analysis of SEM and PROCESS: use of PLS-SEM for mediation analyses! *Int. J. Mark. Res.* 62, 288–299. doi: 10.1177/1470785320915

Sarstedt, M., Ringle, C. M., and Hair, J. F. (2022). "Partial least squares structural equation modeling" in Handbook of market research. eds. C. Homburg, M. Klarmann and A. Vomberg (Cham: Springer International Publishing), 587–632.

Senoner, J., Schallmoser, S., Kratzwald, B., Feuerriegel, S., and Netland, T. (2024). Explainable AI improves task performance in human–AI collaboration. *Sci. Rep.* 14:31150. doi: 10.1038/s41598-024-82501-9

Shmueli, G., Sarstedt, M., Hair, J. F., Cheah, J.-H., Ting, H., Vaithilingam, S., et al. (2019). Predictive model assessment in PLS-SEM: guidelines for using PLSpredict. *Eur. J. Mark.* 53, 2322–2347. doi: 10.1108/ejm-02-2019-0189

Sokha, K. (2024). Examine the impact of contextual, personal, and Behavioral factors on high school teachers' engagement in teaching science using an integrated STEM approach. *Int. J. Sci. Math. Educ.* 22, 1857–1883. doi: 10.1007/s10763-024-10447-5

Sperling, K., Stenberg, C.-J., McGrath, C., Åkerfeldt, A., Heintz, F., and Stenliden, L. (2024). In search of artificial intelligence (AI) literacy in teacher education: a scoping review. *Comp. Educ. Open* 6:100169. doi: 10.1016/j.caeo.2024.100169

Teng, Y., Pu, R., and Hao, Y. (2024). How can an innovative teaching force be developed? The empowering role of distributed leadership. *Think. Skills Creat.* 51:101464. doi: 10.1016/j.tsc.2024.101464

Thanki, H., Tripathy, N., and Shah, S. (2024). Investors' behavioral intention in mutual fund investments in India: applicability of theory of planned behavior. *Asia-Pacific Finan. Markets.* 32, 975–996. doi: 10.1007/s10690-024-09477-4

To, J., Tan, K., and Lim, M. (2023). From error-focused to learner-centred feedback practices: unpacking the development of teacher feedback literacy. *Teach. Teach. Educ.* 131:104185. doi: 10.1016/j.tate.2023.104185

Tzirides, A. O., Zapata, G., Kastania, N. P., Saini, A. K., Castro, V., Ismael, S. A., et al. (2024). Combining human and artificial intelligence for enhanced AI literacy in higher education. *Comp. Educ. Open* 6:100184. doi: 10.1016/j.caeo.2024.100184

Vangrieken, K., and Kyndt, E. (2020). The teacher as an island? A mixed method study on the relationship between autonomy and collaboration. *Eur. J. Psychol. Educ.* 35, 177–204. doi: 10.1007/s10212-019-00420-0

Viberg, O., Cukurova, M., Feldman-Maggor, Y., Alexandron, G., Shirai, S., Kanemune, S., et al. (2024). What explains teachers' trust in AI in education across six countries? *Int. J. Artif. Intell. Educ.* 35, 1288–1316. doi: 10.1007/s40593-024-00433-x

Wahid, M., and Ayub, N. (2024). Predictive role of psychological capital and perceived organizational support on innovative work behavior among higher education teachers of Pakistan. *Tuning J. Higher Educ.* 11, 191–219. doi: 10.18543/tjhe.2715

Wan, C. L. J., Cheung, K., Tsang, H., and Yip, T. L. (2024). Exploring the explanatory variables of peer support: perspectives from direct-entry and transfer students. *Educ. Inf. Technol.* 29, 11445–11481. doi: 10.1007/s10639-023-12162-0

Wang, B. C., Rau, P. L. P., and Yuan, T. Y. (2023). Measuring user competence in using artificial intelligence: validity and reliability of artificial intelligence literacy scale. *Behav. Inform. Technol.* 42, 1324–1337. doi: 10.1080/0144929x.2022.2072768

Wang, Y., and Wang, H. (2024). Mediating effects of artificial intelligence on the relationship between academic engagement and mental health among Chinese college students [original research]. *Front. Psychol.* 15:1477470. doi: 10.3389/fpsyg.2024.1477470

Wang, Q., and Zhao, G. (2021). ICT self-efficacy mediates most effects of university ICT support on preservice teachers' TPACK: evidence from three normal universities in China. *Br. J. Educ. Technol.* 52, 2319–2339. doi: 10.1111/bjet.13141

Wang, X., Zhao, S., Xu, X., Zhang, H., and Lei, V. N.-L. (2025). AI adoption in Chinese universities: insights, challenges, and opportunities from academic leaders. *Acta Psychol.* 258:105160. doi: 10.1016/j.actpsy.2025.105160

Wu, D., Zhou, C., Liang, X., Li, Y., and Chen, M. (2022). Integrating technology into teaching: factors influencing rural teachers' innovative behavior. *Educ. Inf. Technol.* 27, 5325–5348. doi: 10.1007/s10639-021-10815-6

Xiang, B. W., Xin, M. J., Fan, X. D., and Xin, Z. Y. (2024). How does career calling influence teacher innovation? The chain mediation roles of organizational identification and work engagement. *Psychol. Sch.* 61, 4672–4687. doi: 10.1002/pits.23302

Yang, D., Liu, J., Wang, H., Chen, P., Wang, C., and Metwally, A. H. S. (2025). Technostress among teachers: a systematic literature review and future research agenda. *Comput. Hum. Behav.* 168:108619. doi: 10.1016/j.chb.2025.108619

Zhan, Y., Sun, D., Kong, H. M., and Zeng, Y. (2024). Primary school teachers' classroom-based e-assessment practices: insights from the theory of planned behaviour. *Br. J. Educ. Technol.* 55, 2740–2759. doi: 10.1111/bjet.13478

Zhang, Z. (2023). The impact of the artificial intelligence industry on the number and structure of employments in the digital economy environment. *Technol. Forecast. Soc. Chang.* 197:122881. doi: 10.1016/j.techfore.2023.122881

Zhang, H. (2025). Investigating Chinese English learners' readiness for artificial intelligence (AI) technologies: a theory of planned behavior (TPB) perspective. *Learn. Motiv.* 91:102164. doi: 10.1016/j.lmot.2025.102164

Zhang, Y. B., Li, J. F., and Song, J. Y. (2025). The influencing factors of the participation intention and behavior in organized scientific research of young teachers in modern industrial colleges-based on the theory of planned behavior and normative activation theory. *Front. Psychol.* 16:1450240. doi: 10.3389/fpsyg.2025.1450240

Zhao, W., Liu, X., Shah, S., Baah, I., Patel, A., and Wise, N. (2021). Peer support in smart learning and education. 2021 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/IOP/SCI), 598–605.

Zhao, J., and Qin, Y. (2021). Perceived teacher autonomy support and students' deep learning: the mediating role of self-efficacy and the moderating role of perceived peer support. *Front. Psychol.* 12:652796. doi: 10.3389/fpsyg.2021.652796

Zhao, L., Zhang, K., Gao, Y., Jia, Z., and Han, S. (2022). The relationship between gender, marital status and depression among Chinese middle-aged and older people: mediation by subjective well-being and moderation by degree of digitization [original research]. *Front. Psychol.* 13:923597. doi: 10.3389/fpsyg.2022.923597

Zhao, Y., Zhao, K., Chen, I. H., Wang, X., and Wei, S. (2024). Peer support and novice teachers' willingness to participate in teaching reflection in online workshops: the mediating role of perceived competence and sense of belonging. *Curr. Psychol.* 43, 32324–32335. doi: 10.1007/s12144-024-06772-1

Zhou, W., Nakatsubo, F., Wu, J., and Liu, K. (2022). Digital ethnography of an online professional learning community based on WeChat for Chinese early childhood teachers. *Comput. Educ.* 191:104617. doi: 10.1016/j.compedu.2022.104617

Zhou, Y., Thurasamy, R., Yusof, R., Zhang, P., Li, X., and Ling, S. (2025). Factors influencing innovative work behavior among teachers in the higher education sectors in China: the role of work engagement as a mediator and artificial intelligence as a moderator. *Acta Psychol.* 258:105232. doi: 10.1016/j.actpsy.2025.105232

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 24 October 2025
DOI 10.3389/fpsyg.2025.1668633

OPEN ACCESS

EDITED BY
Ivan Wen,
University of Hawaii at Manoa, United States

REVIEWED BY
Gagan Jain,
Manipal University, India
Hongyi Qin,
Zhejiang University, China

*CORRESPONDENCE
Shujun Luo
✉ luoshujun@jxufe.edu.cn

RECEIVED 18 July 2025
ACCEPTED 09 October 2025
PUBLISHED 24 October 2025

CITATION
Chen Y, Luo S and Yin Y (2025) Are you willing to forgive generative AI doctors? Trust repair after failures in online health consultation services.
Front. Psychol. 16:1668633.
doi: 10.3389/fpsyg.2025.1668633

Are you willing to forgive generative AI doctors? Trust repair after failures in online health consultation services

Yanjie Chen[1,2], Shujun Luo[3]* and Yukun Yin[1]

[1]Faculty of Humanities and Arts, Macau University of Science and Technology, Macau, China, [2]Media Art Research Center, Jiangxi Institute of Fashion Technology, Nanchang, China, [3]School of Sociology and Humanities, Jiangxi University of Finance and Economics, Nanchang, China

While generative AI doctors are increasingly used in online health consultation services, research on trust repair following service failures remains limited. We examined how attribution style, social support, and anthropomorphism influence individuals' trust repair and behavioral intention. A total of 512 participants were recruited to take part in a between-subjects experiment with a 2 (internal vs. external attribution) × 2 (informational vs. emotional support) × 2 (anthropomorphism vs. non-anthropomorphism) design. The results revealed that participants exposed to internal attribution, emotional support, or anthropomorphism conditions reported higher levels of trust repair. Anthropomorphism influences the effectiveness of attribution style and social support in repairing trust in GAI doctors. Moreover, an interesting interaction was observed between attribution style and social support: when the GAI doctor used internal attribution, informational support was more effective; under external attribution, emotional support proved more effective. In addition, the effect of social support on behavioral intention was fully mediated by trust repair. These findings offer practical implications for optimizing the design of GAI doctors, enhancing communication and collaboration between GAI doctors and users, and ultimately strengthening the resilience of AI-based health consultation services.

KEYWORDS

generative artificial intelligence (GAI), attribution style, social support, anthropomorphism, trust repair, online health consultation services (OHCSV)

Introduction

In recent years, generative artificial intelligence doctors (GAI doctors) have emerged as a new form of medical assistance and are being widely adopted in online health consultation services (OHCSV; Guo and Chen, 2025; Li, Y et al., 2025). Powered by advanced algorithms, GAI doctors are capable of producing predetermined responses through the analysis of user inputs and retrieval of relevant medical knowledge (Chow et al., 2024). Therefore, compared with human doctors, GAI doctors can provide round-the-clock services, overcome geographical limitations, and supplement scarce medical resources. However, realization of GAI doctors' potential relies heavily on user trust, and low trust or any breach of trust may undermine users' continued engagement with these systems (Li and Liu, 2025; Li, Y et al., 2025). Consequently, many previous studies have focused on how to establish and enhance individuals' trust in GAI doctors (Chen and Cui, 2025; Detjen et al., 2025; Kim et al., 2024). Nevertheless, these studies have mainly addressed the development of general trust, paying little attention to trust repair following service failures. Like any other AI service, GAI doctors are not perfect (Chen et al., 2022). They might also fail, such as providing inaccurate diagnoses, failing to detect important symptoms, or providing suboptimal recommendations with insufficient information. However, unlike other general AI service failures, failures in GAI doctors might cause significant health issues so that people use GAI doctors to seek health care with caution and scrutiny (Quinn et al., 2021). That is, service failures by GAI doctors may notably weaken users' trust and reduce their intention to keep using such services. Hence, focusing on trust repair following service failures of GAI doctors is both practically and theoretically important.

Existing research in the field of human–machine interaction (HMI) indicates that the way trustees attribute the causes of failures significantly influences the trustor's perception of the event (Chen et al., 2022; Kim and Song, 2021). Providing social support by GAI doctors helps enhance individuals' positive expectations toward them (Li et al., 2025; Zhou and Chang, 2024). Endowing GAI doctors with human-like characteristics can improve the resilience of users' trust (De Visser et al., 2016; Li et al., 2023). Despite considerable research on attribution style and anthropomorphism in trust repair, little is known about how these factors affect trust restoration in health consultation scenarios involving GAI doctors. Different forms of social support have been found to affect trust in GAI doctors, but they have seldom been studied in the context of repairing trust after failures. The advancement of medical AI should emphasize human-centered design and trustworthiness (Albahri et al., 2023). In line with this, the present study primarily examines how attribution style, social support, and anthropomorphism influence trust repair in the context of medical AI service failures. In addition, we investigate how trust repair shapes the relationship between social support and behavioral intentions. Gaining insight into these processes can enhance the adaptability and resilience of GAI-based health consultation systems.

Trust and trust repair

In the context of HMI, trust can be defined as the belief or attitude that an agent will assist in achieving an individual's goals in situations characterized by uncertainty and vulnerability (De Visser et al., 2016). Although many scholars define trust and use it as a baseline to study repair, general trust and trust repair differ both qualitatively and quantitatively. From a qualitative perspective, general trust develops under the assumption of "trustworthy until proven otherwise," whereas trust repair occurs after this assumption is violated, with betrayal not only damaging prior trust but also triggering negative emotions and concerns about further harm (Kim et al., 2004; Sharma et al., 2023). Thus, while the essence of general trust lies in fostering positive expectations, trust repair additionally requires addressing post-violation negative effects to restore the relationship. From a quantitative perspective, in the initial stage of a relationship, individuals often exhibit relatively high levels of trust based on cues such as trust propensity, sense of dependence, institutional safeguards, and group identity or reputation (Kim et al., 2004, 2009). However, once a violation occurs, trust can easily fall below its initial level, and the magnitude of increase required to rebuild trust is substantially greater than that needed to establish initial trust (Kim et al., 2004, 2006; Lewicki and Brinsfield, 2017). In summary, trust repair is more complex and challenging than the initial development of general trust. Therefore, this study adopts the definition by Sharma et al. (2023), which states that "trust repair was any increase in trust above the post-transgression level and complete repair as an increase in trust to the pre-transgression level." This definition not only captures the dynamic changes in trust following a violation but also provides a clear operational standard for empirical analysis.

For many years, researchers have focused on exploring the factors and mechanisms that affect trust repair. In general, mechanisms for trust repair can be categorized into attribution, social-equilibrium, and structural mechanisms (Dirks et al., 2009; Sharma et al., 2023). According to attribution mechanisms, after a trust violation occurs, how the trustor attributes the failure plays a major role in restoring the relationship with the trustee (Kim et al., 2009; Tomlinson and Mayer, 2009). Social equilibrium mechanisms suggest that a trust violation disrupts the trust established between parties based on existing social norms, requiring restorative measures, particularly those aimed at alleviating negative emotions, to repair the relationship (Gillespie and Siebert, 2018; Ren and Gray, 2009). Structural mechanisms posit that if the external environment facilitates trust or reduces the likelihood of untrustworthy behaviors, trust can be more effectively restored (Dirks et al., 2009; Sitkin & Roth, 1993). Overall, trust repair primarily involves three dimensions: attribution of the breach, the relationship, and the environment (Sharma et al., 2023). Trust is more likely to be repaired if individuals perceive the attribution of responsibility as acceptable, the damaged relationship is mended, and the environment supports trust. Therefore, based on these three mechanisms, this study aims to examine how attribution style, social support, and anthropomorphism influence trust repair and behavioral intentions in GAI doctors (see Figure 1).

Attribution theory and trust repair

According to attribution theory, attribution constitutes a fundamental cognitive process (Chen et al., 2022; Weiner, 1985). Through this process, individuals seek to identify the causes of behavioral events in order to enhance their understanding of the internal and external world. In general, attributions can be divided into internal and external types. In the context of service failures in HMI, it typically represents different ways of taking responsibility. Specifically, internal attribution means that the GAI takes active responsibility for a service failure, such as attributing it to the use of inaccurate data (Kim and Song, 2021). Conversely, external attribution occurs when the GAI places the cause of a service failure on external factors, such as environmental conditions or human interference (Zhang et al., 2023). Based on expectation confirmation theory, when the attribution style used by a GAI matches individuals' expectations, it is more likely to satisfy their psychological needs and facilitate trust repair (Oliver, 1980). If the attribution style does not match expectations, it could make the negative effects even worse. Studies have shown that following a trust violation, a machine taking responsibility proactively helps repair trust because it signals sincere regret (Kim et al., 2006; Ohbuchi et al., 1989; Tomlinson et al., 2004). However, some studies suggest that proactively taking responsibility does not always produce positive outcomes. For example, Kim and Song (2021) found that when an anthropomorphized AI issued an apology based on external rather than internal attribution, it resulted in greater trust repair. Furthermore, some researchers have found that internal attribution tends to elicit blame from the victim, whereas external attribution does not, as people

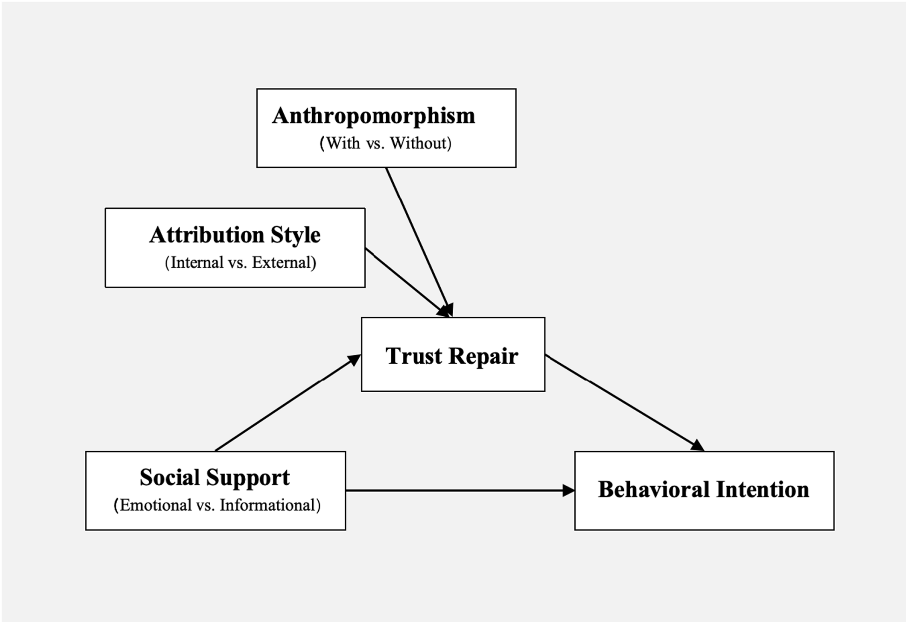


FIGURE 1
Conceptual model in the current study.

recognize that many events are influenced by external factors (Kim et al., 2006; Sullivan, 1975). Yet, external attributions are not without drawbacks. When trustors question the agent's innocence, such attributions may be perceived as excuses or indications of incompetence (Schlenker et al., 2001). Kim et al. (2006) found that in human-to-human interaction (HHI), internal attributions for competence-related failures are more effective than external attributions in repairing trust, as they convey responsibility and integrity to the trustor and, more importantly, signal a greater likelihood of correcting the behavior in the future. GAI, supported by large-scale machine learning models, can continuously optimize its algorithms through iterative training, thereby enhancing the quality and adaptability of its outputs (Qin et al., 2025). Therefore, in the context of this study, we propose the following hypothesis:

> *H*1: Compared with external attributions, internal attributions will result in higher trust repair.

Social support and trust repair

Prior studies have shown that trust violations in HMI may be alleviated through trust repair strategies, such as the provision of recovery services (Kim and Song, 2021; Meng et al., 2025). More specifically, service recovery involves the actions a provider implements following a service failure, aimed at mitigating customer dissatisfaction and resolving complaints—typically through apology, compensation, and restoration (Spreng et al., 1995; Zhou and Chang, 2024). According to social support theory, individuals' access to supportive relationships or resources—primarily in the form of informational support and emotional support—can have a positive impact on their well-being (Langford et al., 1997). Informational support means offering useful guidance and advice to assist individuals in solving problems and making informed decisions (Madjar, 2008). Emotional support involves the expression of love, empathy, and understanding, allowing individuals to feel cared for and understood (Reblin and Uchino, 2008). Accordingly, social support theory has been extensively used in trust-building research. However, few studies have examined how social support influences individuals' trust repair, particularly in the context of AI-based health consultations. Specifically, in the domain of OHCSV, GAI doctors can provide informational service recovery by explaining the reasons for service failures and offering additional informational support to help individuals address their concerns (Zhou and Chang, 2024). Previous research indicates that due to the black-box nature of AI, lay users often lack understanding of how decisions or results are generated. Therefore, informing users about the AI system's data processing and operational mechanisms is considered an effective approach to enhancing user trust (Afroogh et al., 2024; Felzmann et al., 2019). In other words, a substantial body of prior research has demonstrated that the provision of transparent information helps users feel neither deceived nor compelled. However, numerous studies have also demonstrated that trust is not a simple function of transparency; human-like features of robots, particularly emotional attributes, play a significant role in facilitating interaction between humans and AI (Gebhard et al., 2021; Troshani et al., 2021). Emotional service recovery can allow individuals to feel understood, empathized with, and comforted by the AI, thereby potentially alleviating the negative experiences caused by service failures. Given that, in the context of service failures during health consultations, individuals primarily experience pressure to obtain clear, accurate, and useful medical information to reduce uncertainty and guide their health decisions (Li, Y et al., 2025; Liu et al., 2022), we predict that informational support, compared with emotional support, will be more effective in facilitating trust repair.

> *H*2: Compared with emotional support, informational support will result in higher trust repair.

Moreover, previous research has consistently shown that social support positively affects users' behavioral intentions (Bu et al., 2024; Rashidi et al., 2025; Zhou and Chang, 2024); yet, service failures may weaken this effect, reducing continued engagement with GAI healthcare services. Trust is crucial in designing interactive intelligent agents, as it influences how individuals perceive, interact with, and evaluate technology (Kim and Song, 2021; Li et al., 2008). Based on this, we argue that in the context of GAI doctor service failures, trust repair may play a key role in the relationship between social support and behavioral intention. Accordingly, the following hypotheses are proposed:

> *H*3a: Social support positively influences behavioral intention.
>
> *H*3b: Trust repair mediates the relationship between social support and behavioral intention.

Anthropomorphism and trust repair

With the rapid advancement of technologies such as robotics, automation, and natural language processing, the boundary between humans and machines has become increasingly blurred (De Visser et al., 2016). Robots are not only becoming more intelligent and capable of assisting humans across various domains, but are also increasingly anthropomorphized, as designers often incorporate human-like visual features, identity cues, or language to enhance their social presence (Go and Sundar, 2019). According to the Computers Are Social Actors (CASA) paradigm, enhancing the level of anthropomorphism in machines facilitates HMI by making the agent appear more familiar and trustworthy (Nass et al., 1994). In service recovery contexts, existing research similarly suggests that anthropomorphism improves consumer experience and enhances the effectiveness of service recovery. For example, Agnihotri and Bhattacharya (2024) demonstrated that anthropomorphism enhances consumers' perceptions of a chatbot's honesty and integrity, thereby increasing their willingness to forgive it for service failures. Zhou and Chang (2024) reported a positive association between higher levels of anthropomorphism and both perceived service quality and attitude satisfaction in service recovery contexts. Moreover, De Visser et al. (2016) found that anthropomorphism enhances trust resilience in cognitive agents. Although anthropomorphism's positive effects on service recovery have been widely studied, its role in trust repair specifically within AI healthcare consultations receives limited attention. Li, Y et al. (2025) showed that in AI healthcare consultations, anthropomorphism boosts perceptions of a robot's social presence, increasing source credibility and behavioral intentions. This suggests people apply different "humanness" heuristics when interacting with robots versus real humans, resulting in distinct psychological responses (Li, Y et al., 2025; Sundar, 2008). Based on this, the current study assumes that anthropomorphism also improves the effectiveness of trust repair in AI healthcare consultations. Accordingly, we propose the following research hypothesis:

> *H*4: Compared with non-anthropomorphic GAI doctors, anthropomorphic GAI doctors will result in higher trust repair.

In addition to examining the main effects of attribution style, social support, and anthropomorphism on trust repair, this study also explores whether there are interaction effects among these factors. According to Kim and Song (2021), the lowest level of trust damage occurred when a machine-like agent used external rather than internal attributions. Li, Y et al. (2025) reported that anthropomorphic GAI doctors providing informational support can enhance their social presence, thereby increasing source credibility. Moreover, Chen et al. (2022) found that in cases of service failure with external attribution, recovery actions taken by the healthcare provider, rather than the consumer, were effective in restoring cognitive trust. Therefore, we hypothesize that attribution style, social support, and anthropomorphism interactively affect trust repair in GAI doctors:

> *H*5: There is an interaction effect between attribution style, social support, and anthropomorphism on trust repair.

Methods

Participants

This study recruited 512 eligible participants through Credamo, an online experimental survey platform specializing in social science research in China. All participants were over 18 years old and met the inclusion criteria (see Table 1). They were randomly selected from Credamo's managed respondent pool. We performed *a priori* power analysis with G*Power 3.1 software to confirm sufficient statistical power. The results presented that at least 210 participants were needed (power = 0.95, α = 0.05, effect size = 0.25), a requirement that our sample successfully fulfilled.

Design

Upon the approval of IRB of the author's affiliated university (MUST-FA-20250017), we conducted an online experiment with a 2

TABLE 1 Demographic characteristics of participants.

Demographics variable	Category	Frequency
Gender	Female	356
	Male	156
Age	< 20	4
	20–29	281
	30–39	189
	40–49	23
	50–59	14
	60+	1
Education	junior college or below	42
	Undergraduate	375
	Master's degree and above	95
Frequency of using GAI doctors	< 5 times	1
	5–10 times	121
	11–15 times	263
	16–20 times	109
	> 20 times	18

(internal attribution vs. external attribution) × 2 (informational support vs. emotional support) × 2 (anthropomorphism vs. non-anthropomorphism) between-subjects factorial design. Two medical professionals were invited to review the AI-generated content for accuracy.

The experiment included two scenarios and three stages of trust measurement: initial trust, trust violation, and trust repair. Scenario 1 (Trust Violation) presented a text-only dialog in which the GAI doctor's advice conflicted with participants' prior knowledge, aiming to induce a decline in trust. Scenario 2 (Trust Repair) built upon Scenario 1, presenting the full dialog including the trust violation and the assigned recovery strategy, in order to examine how different combinations of attribution style, social support, and anthropomorphism influenced trust repair (see Supplementary materials). Notably, Scenario 1 constituted the first part of Scenario 2, since trust repair logically requires a prior violation. To prevent the manipulation of anthropomorphism from influencing the trust violation scenario, Scenario 1 was presented in a text-only format.

At the beginning of the experiment, participants reported their initial trust in the GAI doctor after providing informed consent, serving as a baseline measurement. Next, participants entered Scenario 1, where they were asked to imagine consulting the GAI doctor about fish oil consumption (viewing the stimulus for at least 15 s) and then report their trust in the GAI doctor. Subsequently, participants were randomly assigned to one of the eight experimental conditions (Scenario 2). During this scenario, participants viewed the full dialog between the GAI doctor and the patient (for at least 35 s) and then reported their trust in the doctor again. Additionally, participants reported their behavioral intentions and demographic information, including gender, age, education, and frequency of using GAI doctors. Finally, participants were explicitly informed that the information provided was fictitious and did not constitute real medical advice.

Stimulus

For this study, the experimental dialog was set within a scenario in which users inquired about the appropriate dosage of fish oil supplements. This scenario was chosen due to the growing attention individuals pay to personal health management. Although people frequently purchase dietary supplements independently, they often lack sufficient knowledge regarding their necessity and correct usage. Within this health-consumption context, consulting GAI doctors has become a convenient way for individuals to access health advice.

Following previous research (Kim and Song, 2021), we manipulated attribution style by defining internal attribution as errors in AI health consultations caused by the system retrieving inaccurate information, and external attribution as errors resulting from insufficient information provided by the user. Accordingly, participants in the internal attribution condition were presented with a GAI doctor attributing the error to the AI system itself, whereas those in the external attribution condition saw the GAI doctor attributing the error to the user.

For social support, participants in the informational support condition were exposed to a GAI doctor that appeared objective and calm, offering detailed advice on fish oil supplementation. Example expressions included specific dosage recommendations such as, "Relevant studies suggest that a daily intake of 1,000 to 3,000 mg of fish oil is generally safe and beneficial for healthy adults," along with links to additional web resources for further information. In the emotional support condition, participants were exposed to a GAI doctor conveying warmth and understanding. Example expressions included, "Dear friend, I truly understand your concern about your health, and I know how confusing it could be when faced with so much conflicting information. I'll always be here with you, supporting and protecting your health."

Moreover, we adopted the approach of manipulating anthropomorphic visual cues based on prior research (Go and Sundar, 2019; Li, Y et al., 2025). For participants in the anthropomorphism condition, the interaction interface featured a fictional GAI doctor with human-like characteristics. In contrast, those in the non-anthropomorphism condition viewed a standard ChatGPT dialog window.

Measures

Trust repair

A three-item scale adapted from Meng et al. (2025) was used to measure trust repair, with participants rating each item on a 7-point Likert scale (1 = strongly disagree, 7 = strongly agree). The items were: (1) The GAI doctor gives me the impression of being trustworthy; (2) I consider the GAI doctor to be competent and reliable; (3) I think GAI doctors are willing to look after the health interests of patients (M = 3.876, SD = 1.854, *Cronbach's* α = 0.894). Trust at the initial, trust violation, and trust repair stages was measured using the same scale.

Behavioral intention

A four-item scale adapted from Hadi et al. (2024) was used to measure behavioral intention, with participants rating each item on a 7-point Likert scale (1 = strongly disagree, 7 = strongly agree). The items were: (1) I intend to continue using AI health consultation; (2) Compared to other consultation methods, I am still willing to consult a GAI doctor; (3) I am willing to consult a GAI doctor again when I face health issues in the future; (4) It is unlikely that I will stop using AI health consultation because of a service failure problem (M = 4.254, SD = 2.499, *Cronbach's* α = 0.941).

To assess the effectiveness of our experimental manipulations, we included three sets of manipulation check items in the questionnaire. For attribution, participants were invited to answer the question: "Was the service failure caused by the AI system retrieving inaccurate information?" To evaluate social support, participants rated the GAI doctor on perceived sympathy, inspiration, warmth, and care. Higher scores indicated a greater level of emotional support. For anthropomorphism, participants answered the question: "How do you think about the GAI doctor's anthropomorphism capability?" A 7-point Likert scale (1 = strongly disagree, 7 = strongly agree) was used to assess all items.

Results

Data analysis

Since this study involved two scenarios and three stages of trust measurement, a paired-samples t-test was conducted to examine changes in trust across the stages. The results, presented in Table 2, indicate that trust significantly decreased following the service failure and was subsequently restored after recovery, regardless of the recovery method. These findings confirm that the manipulation was successful, allowing us to proceed with further analyses.

Randomization check

To examine whether participants were successfully randomized across conditions, a series of chi-square tests and one-way ANOVAs were conducted. Results showed no significant differences among the eight experimental groups in terms of gender ($\chi^2(7) = 5.695$, $p = 0.576$), age ($F(7, 504) = 1.557$, $p = 0.146$), education ($F(7, 504) = 1.054$, $p = 0.393$), or frequency of using GAI doctors ($F(7, 504) = 0.348$, $p = 0.932$).

Manipulation check

Given the $2 \times 2 \times 2$ between-subjects design, t-tests for independent groups were conducted to assess the effectiveness of the manipulations of attribution style, social support and anthropomorphism (see Table 3). Results confirmed the success of the manipulations. Participants exposed to internal attribution ($M = 6.287$, $SD = 0.785$) conveyed significantly stronger perceptions of internal attribution than those exposed to external attribution ($M = 3.543$, $SD = 1.853$), $t(510) = 21.885$, $p < 0.001$. Similarly, participants assigned to the emotional support condition ($M = 5.053$, $SD = 1.095$) perceived significantly greater emotional support compared to those in the informational support condition ($M = 3.543$, $SD = 1.238$), $t(510) = 14.611$, $p < 0.001$. Moreover, significantly higher perceived anthropomorphism was reported by participants in the anthropomorphic condition ($M = 4.713$, $SD = 1.111$) than those in the non-anthropomorphic condition ($M = 3.977$, $SD = 1.200$), $t(510) = 7.204$, $p < 0.001$.

TABLE 2 The comparison among the trust in three stages.

Outcome variable	Stage	*M*	*SD*	*t*-value
Trust	Initial-violation	2.159	1.222	39.986***
	Violation-repaired	−0.687	1.300	−11.958***

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

Main findings

Hypothesis testing

A three-way analysis of variance (ANOVA) was conducted with attribution style, social support, and anthropomorphism as independent variables and trust repair as the dependent variable (see Table 4). The results revealed significant main effects of attribution style, social support, and anthropomorphism on trust repair. Regarding attribution style, participants in the internal attribution condition showed greater trust repair ($M = 3.987$, $SD = 1.150$) compared to those in the external attribution condition ($M = 3.766$, $SD = 1.309$), $F(1, 504) = 4.183$, $p < 0.05$. For social support, participants in the emotional support condition reported higher trust repair ($M = 3.983$, $SD = 1.268$) than those in the informational support condition ($M = 3.766$, $SD = 1.196$), $F(1, 504) = 4.118$, $p < 0.05$. In addition, participants exposed to the anthropomorphic condition reported higher trust repair ($M = 4.033$, $SD = 1.186$) than those in the non-anthropomorphic condition ($M = 3.721$, $SD = 1.267$), $F(1, 504) = 8.247$, $p < 0.01$. Thus, H1 and H4 were supported, while H2 was not.

Regarding H5, significant interaction effects on trust repair were found for the interactions between anthropomorphism and attribution style ($F(1, 504) = 5.994$, $p < 0.05$), anthropomorphism and social support ($F(1, 504) = 4.724$, $p < 0.05$), and attribution style and social support ($F(1, 504) = 4.947$, $p < 0.05$). Regarding the interaction between anthropomorphism and attribution style, Figure 2 presents a plot of the obtained mean scores. In the anthropomorphic condition, external attribution was more effective in repairing trust, whereas in the non-anthropomorphic condition, internal attribution was more effective. Specifically, individuals who were assigned to the anthropomorphic-external attribution condition reported higher trust repair ($M = 4.055$, $SD = 1.256$) than those in the anthropomorphic-internal attribution condition ($M = 4.010$, $SD = 1.117$), the non-anthropomorphic-internal attribution condition ($M = 3.963$, $SD = 1.187$), and the non-anthropomorphic-external attribution condition ($M = 3.486$, $SD = 1.303$). A similar pattern emerged for the interaction between anthropomorphism and social support. As shown in Figure 3, individuals in the anthropomorphic–emotional support condition reported higher trust repair ($M = 4.255$, $SD = 1.160$) than those in the anthropomorphic–informational support condition ($M = 3.807$, $SD = 1.175$), the non-anthropomorphic–informational support condition ($M = 3.727$, $SD = 1.221$), and the non-anthropomorphic–emotional support condition ($M = 3.715$, $SD = 1.317$), indicating that trust repair is greatest when information combines anthropomorphism with emotional support. As for the interaction between attribution style and social support, Figure 4

TABLE 3 T-test of experimental manipulation.

Group	Number	*M*	*SD*	*t*	*df*	*p*
Internal Attribution	254	6.287	0.785	21.885	510	0.001
External Attribution	258	3.543	1.853			
Informational Support	254	3.543	1.238	14.611	510	0.001
Emotional Support	258	5.053	1.095			
Anthropomorphism	254	4.713	1.111	7.204	510	0.001
Non-Anthropomorphism	258	3.977	1.200			

TABLE 4 Attribution style x social support x anthropomorphism factorial analysis of variance for trust repair.

Source	df	F	η^2	p
Attribution Style	1.000	4.183	0.008	0.041
Social Support	1.000	4.118	0.008	0.043
Anthropomorphism	1.000	8.247	0.016	0.004
Anthropomorphism x Attribution Style	1.000	5.994	0.012	0.015
Anthropomorphism x Social Support	1.000	4.724	0.009	0.030
Attribution Style x Social Support	1.000	4.947	0.010	0.027
Anthropomorphism x Attribution Style x Social Support	1.000	0.080	0.000	0.777
Error	504			

presents the mean scores. When internal attribution was used, informational support was more effective in repairing trust, whereas under external attribution, emotional support led to higher levels of trust repair. Specifically, individuals in the internal attribution–informational support condition reported the highest trust repair ($M = 3.997$, $SD = 1.136$) compared to those in the external attribution–emotional support condition ($M = 3.990$, $SD = 1.363$), the internal attribution–emotional support condition ($M = 3.977$, $SD = 1.168$), and the external attribution–informational support condition ($M = 3.539$, $SD = 1.215$; see Table 5).

In addition, no significant three-way interaction was observed among anthropomorphism, attribution style, and social support on trust repair ($F(1, 504) = 0.080$, $p = 0.777$).

Mediation analysis

The mediating role of trust repair was examined using PROCESS Model 4 with 5,000 bootstrap samples. The results showed that social support significantly predicted trust repair ($b = 0.217$, $SE = 0.109$, $p = 0.047$), and trust repair significantly predicted behavioral intention ($b = 0.899$, $SE = 0.034$, $p < 0.001$). However, the direct effect of social support on behavioral intention was not significant ($b = 0.036$, $SE = 0.084$, $p = 0.671$). Importantly, the indirect effect of social support on behavioral intention via trust repair was significant (indirect effect = 0.195, BootSE = 0.097, 95% CI [0.002, 0.380]; see Figure 5). These findings suggest that trust repair serves as a full mediator between social support and behavioral intention, thus supporting H3b while H3a is not supported.

Discussion

This study was primarily designed to examine trust repair of GAI doctors in the context of online health consultation service failures. Specifically, we investigated the main and interaction effects of attribution style, social support, and anthropomorphism on trust repair, as well as the relationships among social support, trust repair, and behavioral intention.

Firstly, the main effect of attribution style was examined. Results revealed greater trust repair when internal attribution was provided by the GAI doctor compared to external attribution. This may be because when GAI doctors actively take responsibility, individuals may perceive that the GAI doctor has recognized the problem and will take corrective actions, thus fostering positive expectations for the quality of subsequent interactions (Kim et al., 2006). Regarding social support, emotional support proved more effective for trust repair than informational support. A possible explanation is that, following failures in AI-based healthcare services, offering empathy and emotional support may be more critical for individuals than simply providing information. According to Meng and Dai (2021), providing emotional support—whether in HHI or HMI—helps individuals feel supported, thereby alleviating stress and anxiety. Moreover, our study found that anthropomorphism enhances trust repair in AI health consultation failures, consistent with prior research (De Visser et al., 2016; Meng et al., 2025). This suggests that designing GAI doctors with anthropomorphic features to enhance trust resilience is a crucial goal in HMI (De Visser et al., 2016). Considering the current low adoption rates of medical AI, enhancing the social characteristics of GAI doctors may improve public attitudes and increase tolerance for service failures. It is noteworthy that, although attribution style, social support, and anthropomorphism significantly influenced trust repair, trust during the repair stage ($M = 3.876$) was only slightly higher than after the violation ($M = 3.189$) and remained below initial trust ($M = 5.346$). This aligns with previous findings that trust rarely fully recovers after a violation (Kim et al., 2009; Lewicki and Brinsfield, 2017). Our study further indicates that, in the context of health consultations, trust in GAI doctors is particularly difficult to restore.

Secondly, significant interactions were found between anthropomorphism and attribution style, and between anthropomorphism and social support, both revealing a similar pattern: anthropomorphism alters the psychological framework individuals use to evaluate GAI doctors. Specifically, when interacting with an anthropomorphic GAI doctor, individuals are more likely to employ a "human heuristic," perceiving them as social actors with intentions and emotions. In contrast, when interacting with a non-anthropomorphic GAI doctor, individuals tend to adopt a "machine heuristic," viewing them as technical tools devoid of social capabilities (Nass et al., 1994; Sundar, 2008). Therefore, for anthropomorphic GAI doctors, external attribution is more effective in repairing trust, possibly because patients perceive them as "human-like agents" and are thus more likely to understand and forgive their mistakes (De Visser et al., 2016). In contrast, for non-anthropomorphic GAI doctors, internal attribution better facilitates trust repair, aligning with patients' expectations that "technical tools should be responsible and self-correcting" (Coeckelbergh, 2022). Thus, following a trust violation, internal attribution by a non-anthropomorphic GAI doctor appears more sincere and transparent, whereas external attribution may lead patients to perceive a shirking of responsibility, thereby undermining

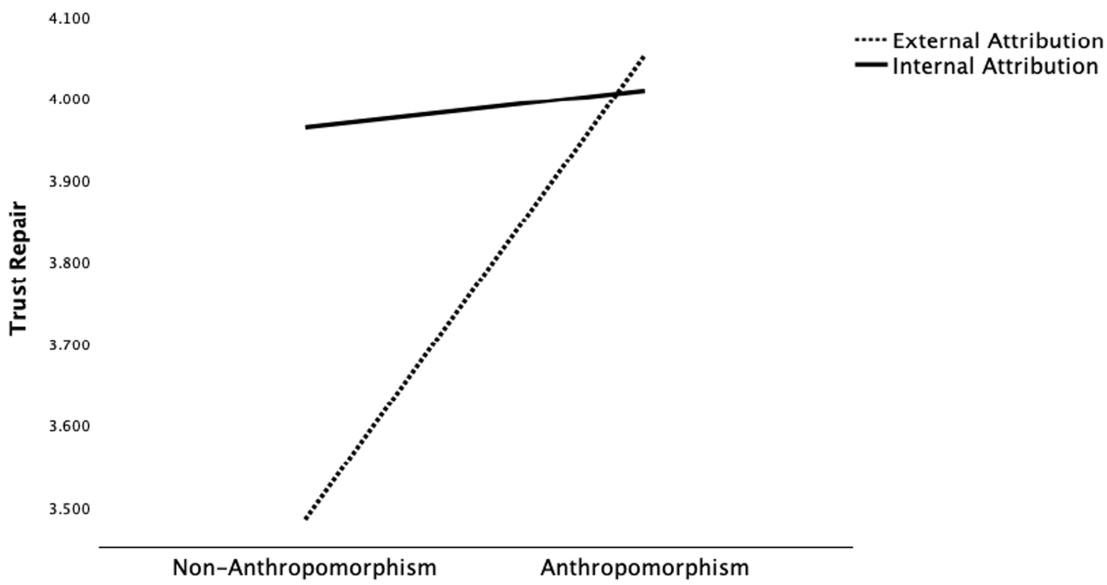


FIGURE 2
Interactive effects between anthropomorphism and attribution style on trust repair.

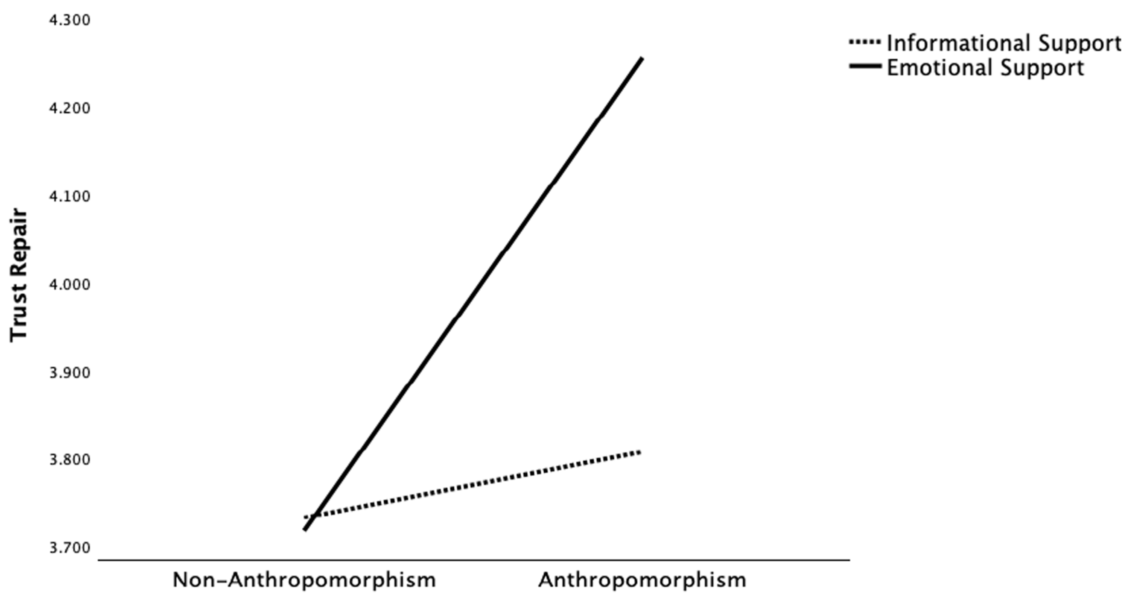


FIGURE 3
Interactive effects between anthropomorphism and social support on trust repair.

trust repair. Similarly, when GAI doctors are anthropomorphic, providing emotional support such as care and reassurance aligns with the human heuristic, making patients perceive them as socially present and sincere, thereby facilitating trust repair more effectively. Meng and Dai (2021) found that the same emotionally supportive messages were perceived as more beneficial when they came from a human partner rather than a chatbot. Overall, the study finds that anthropomorphism influences trust repair by shaping whether individuals adopt a "human heuristic" or a "machine heuristic," which in turn affects the effectiveness of attribution strategies and supportive communication.

In addition, the study also found a significant interaction effect between attribution style and social support. That is, when internal attribution was used, informational support proved to be more effective in repairing trust, and when external attribution was used, emotional support led to better trust repair. This is an interesting result, which indicates that GAI doctors do not always need to take full responsibility for service failures. Instead, they can strategically adjust their support approach based on the type of attribution applied. When the service failure results from external factors, such as the patient providing insufficient information, offering emotional support can help bridge the relational gap between the GAI doctor and the patient. In previous

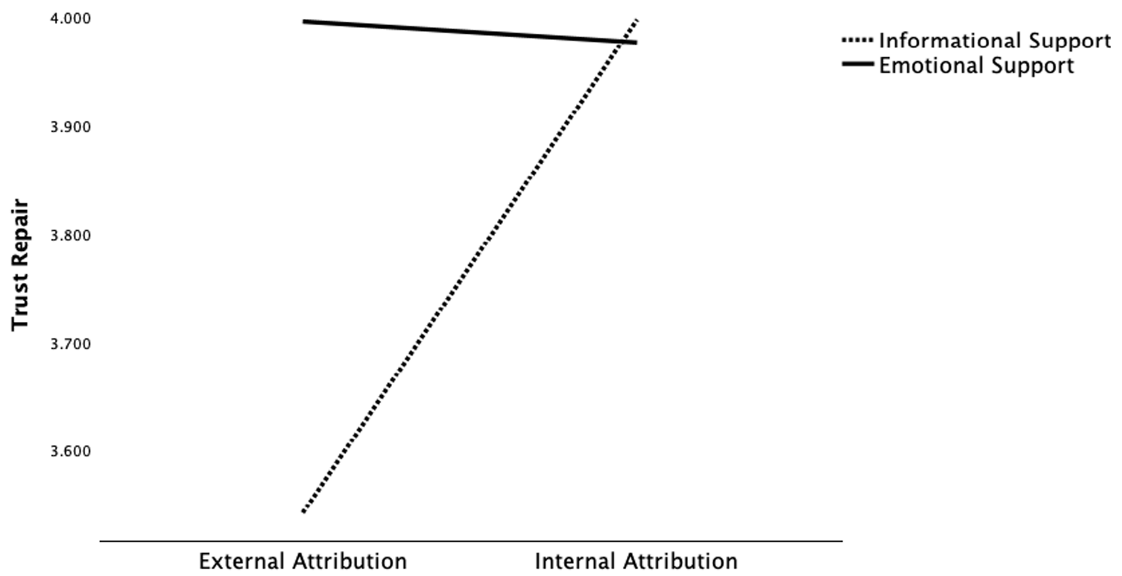


FIGURE 4
Interactive effects between attribution style and social support on trust repair.

TABLE 5 Descriptive statistics for trust repair.

Attribution style	Social support	Anthropomorphism	*N*	Mean	*SD*
External	Informational	Without	65	3.359	1.149
		With	63	3.725	1.261
		Total	128	3.539	1.215
	Emotional	Without	66	3.611	1.437
		With	64	4.380	1.171
		Total	130	3.990	1.363
	Total	Without	131	3.486	1.303
		With	127	4.055	1.256
		Total	258	3.766	1.309
Internal	Informational	Without	63	4.106	1.184
		With	63	3.889	1.086
		Total	126	3.997	1.136
	Emotional	Without	64	3.823	1.182
		With	64	4.130	1.143
		Total	128	3.977	1.168
	Total	Without	127	3.963	1.187
		With	127	4.010	1.117
		Total	254	3.987	1.150
Total	Informational	Without	128	3.727	1.221
		With	126	3.807	1.175
		Total	254	3.766	1.196
	Emotional	Without	130	3.715	1.317
		With	128	4.255	1.160
		Total	258	3.983	1.268
	Total	Without	258	3.721	1.267
		With	254	4.033	1.186
		Total	512	3.876	1.237

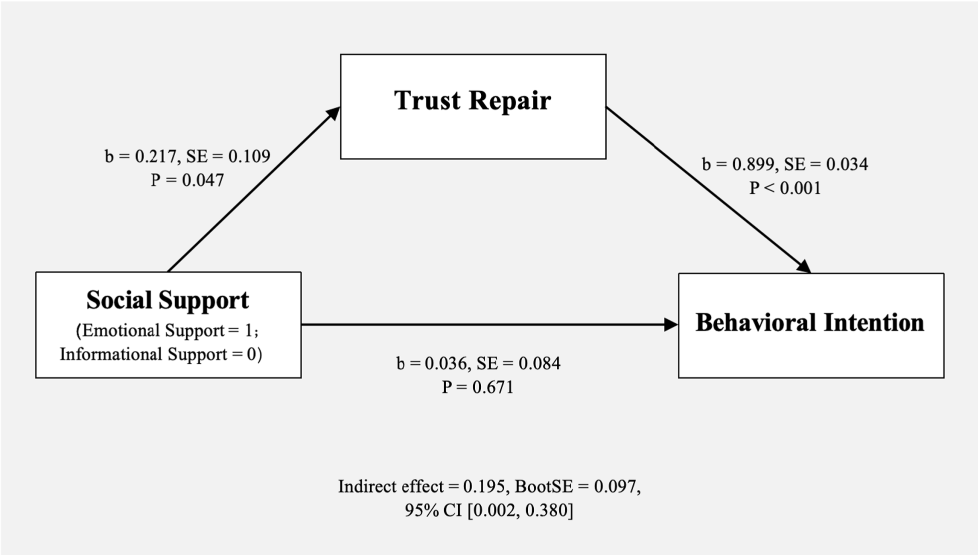


FIGURE 5
Mediation model.

studies, researchers have expressed concerns that when AI frequently makes internal attributions, it may be blamed by participants, whereas when AI makes external attributions, participants are more likely to perceive it as incompetent or making excuses (Kim et al., 2006; Kim and Song, 2021). Our results imply that when external attribution is used, providing emotional support can inherently make individuals feel understood and supported, rather than perceiving the AI as avoiding responsibility. In contrast, when internal attribution is adopted, offering informational support can help individuals better understand the causes behind the GAI doctor's error and receive appropriate solutions, thereby mitigating potential negative effects and facilitating trust repair.

Finally, the study found that social support did not influence behavioral intentions, and trust repair fully mediated this relationship. This result further highlights that the credibility of medical AI plays a decisive role in users' willingness to use its services.

Limitations and implications

Our study has several theoretical contributions. First, since most prior trust repair research has focused on non-health contexts (Kim and Song, 2021; Meng et al., 2025; Wu et al., 2025), investigating GAI doctors contributes to expanding the trust repair literature. Second, previous studies have primarily focused on the effects of attribution style and anthropomorphism on trust repair (De Visser et al., 2016; Zhang et al., 2023), while the role of social support and its interactions with the other two factors in influencing trust repair has been rarely examined. This research offers a comprehensive perspective on how trust can be repaired in interactions with GAI doctors. Additionally, existing research has produced inconsistent findings regarding the effectiveness of different attribution styles on trust repair (Kim et al., 2006; Wu et al., 2025). We found that trust repair is facilitated when internal attribution is paired with informational support and when external attribution is paired with emotional support. These findings make a significant contribution to the body of knowledge on attribution theory.

In terms of practical implications, the interactions between anthropomorphism and attribution style, as well as between anthropomorphism and social support, suggest that trust repair strategies should pay attention to the individual characteristics of GAI doctors. Moreover, the interaction between attribution style and social support indicates that GAI doctors do not always need to assume full responsibility following service failures. Based on the operationalization of external attribution in this study—that service failures result from insufficient information provided by users—this may imply that some medical service failures can be addressed by encouraging users to re-engage in the dialog. This suggests that AI designers could focus on fostering collaborative communication between GAI doctors and users, rather than relying solely on the AI's performance, to more effectively enhance trust repair.

This study has its limitations. Firstly, although the main effects of social support, attribution style, and anthropomorphism on trust repair were statistically significant in this study, the absolute differences between conditions were relatively small. This may be related to the cross-sectional design of the experimental stimuli. Future research could develop simulated online health consultation systems, allowing GAI doctors to engage in multiple rounds of interaction with patients, thereby enabling patients to more clearly perceive the effects of different experimental conditions. Moreover, future studies could explore additional factors that may have a stronger impact on trust repair. Secondly, this study examined trust repair in different stages of GAI doctors' service failures only in an online experiment, without considering longer-term relationships. Future research could adopt a longitudinal design to track users' trust changes following service failures, allowing for a deeper analysis of the trust repair process. Finally, this study did not investigate the influence of individual characteristics on trust repair in AI health consultation service failure contexts. Future research could explore how variables such as AI literacy, previous experience with online medical services, and socioeconomic status affect trust repair.

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 humans were approved by Research Ethics Committee of Faculty of Humanities and Arts at Macau University of Science and Technology. The studies were conducted in accordance with the local legislation and institutional requirements. The 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

YC: Project administration, Conceptualization, Writing – review & editing, Supervision, Formal analysis, Software, Investigation, Data curation, Resources, Writing – original draft, Methodology. SL: Software, Visualization, Resources, Validation, Formal analysis, Writing – original draft, Methodology. YY: Software, Writing – original draft, Supervision.

Funding

The author(s) declare that no financial support was received for the research and/or publication of this article.

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.

Generative AI statement

The author(s) declare that no Gen AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

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/fpsyg.2025.1668633/full#supplementary-material

References

Afroogh, S., Akbari, A., Malone, E., Kargar, M., and Alambeigi, H. (2024). Trust in AI: progress, challenges, and future directions. *Humanit. Soc. Sci. Commun.* 11:1568. doi: 10.1057/s41599-024-04044-8

Agnihotri, A., and Bhattacharya, S. (2024). Chatbots' effectiveness in service recovery. *Int. J. Inf. Manag.* 76:102679. doi: 10.1016/j.ijinfomgt.2023.102679

Albahri, A. S., Duhaim, A. M., Fadhel, M. A., Alnoor, A., Baqer, N. S., Alzubaidi, L., et al. (2023). A systematic review of trustworthy and explainable artificial intelligence in healthcare: assessment of quality, bias risk, and data fusion. *Inf. Fusion* 96, 156–191. doi: 10.1016/j.inffus.2023.03.008

Bu, D., Zhang, C.-Q., Liang, W., Han, Z., Yi, N., Su, N., et al. (2024). Mental health literacy and help-seeking intention among Chinese elite athletes: the mediating roles of stigma and social support. *Front. Psychol.* 15:1332343. doi: 10.3389/fpsyg.2024.1332343

Chen, A., Pan, Y., Li, L., and Yu, Y. (2022). Are you willing to forgive AI? Service recovery from medical AI service failure. *Ind. Manag. Data Syst.* 122, 2540–2557. doi: 10.1108/IMDS-12-2021-0801

Chen, C., and Cui, Z. (2025). Impact of AI-assisted diagnosis on American patients' trust in and intention to seek help from health care professionals: randomized, web-based survey experiment. *J. Med. Internet Res.* 27:e66083. doi: 10.2196/66083

Chow, J. C. L., Wong, V., and Li, K. (2024). Generative pre-trained transformer-empowered healthcare conversations: current trends, challenges, and future directions in large language model-enabled medical chatbots. *BioMedInformatics* 4, 837–852. doi: 10.3390/biomedinformatics4010047

Coeckelbergh, M. (2022). *Self-improvement: Technologies of the soul in the age of artificial intelligence*. New York: Columbia University Press.

Detjen, H. H. J., Densky, L., Von Kalckreuth, N., and Kopka, M. (2025). "Who is trusted for a second opinion? Comparing collective advice from a medical AI and physicians in biopsy decisions after mammography screening." *Proceedings of the 2025 CHI conference on human factors in computing systems*, 1–15.

De Visser, E. J., Monfort, S. S., McKendrick, R., Smith, M. A. B., McKnight, P. E., Krueger, F., et al. (2016). Almost human: anthropomorphism increases trust resilience in cognitive agents. *J. Exp. Psychol. Appl.* 22, 331–349. doi: 10.1037/xap0000092

Dirks, K. T., Lewicki, R. J., and Zaheer, A. (2009). Reparing relationships within and between organizations: building a conceptual foundation. *Acad. Manag. Rev.* 34, 68–84. doi: 10.5465/amr.2009.35713285

Felzmann, H., Villaronga, E. F., Lutz, C., and Tamò-Larrieux, A. (2019). Transparency you can trust: transparency requirements for artificial intelligence between legal norms and contextual concerns. *Big Data Soc.* 6:2053951719860542. doi: 10.1177/2053951719860542

Gebhard, P., Aylett, R., Higashinaka, R., Jokinen, K., Tanaka, H., and Yoshino, K. (2021). "Modeling trust and empathy for socially interactive robots" in *Multimodal agents for ageing and multicultural societies: Communications of NII Shonan meetings*. eds. J. Miehle, W. Minker, E. André and K. Yoshino (Singapore: Springer), 21–60.

Gillespie, N., and Siebert, S. (2018). Organizational trust repair. In R. H. Searle, A. I. Nienaber and S. B. Sitkin (Eds.), *The Routledge companion to trust*. Abingdon, Oxon: Routledge. pp. 284–301.

Go, E., and Sundar, S. S. (2019). Humanizing chatbots: the effects of visual, identity and conversational cues on humanness perceptions. *Comput. Hum. Behav.* 97, 304–316. doi: 10.1016/j.chb.2019.01.020

Guo, W., and Chen, Y. (2025). Investigating whether AI will replace human physicians and understanding the interplay of the source of consultation, health-related stigma, and explanations of diagnoses on patients' evaluations of medical consultations: randomized factorial experiment. *J. Med. Internet Res.* 27:e66760. doi: 10.2196/66760

Hadi, Z. A., Siregar, D. A., Wijaya, G. S. T., Handayani, P. W., and Harahap, N. C. (2024). The influence of transparency, anthropomorphism, and positive politeness on chatbots for service recovery in E-health applications. *Cogent Soc. Sci.* 10:2415534. doi: 10.1080/23311886.2024.2415534

Kim, D., Vegt, N., Visch, V., and Bos-De Vos, M. (2024). How much decision power should (a) I have?: investigating patients' preferences towards AI autonomy in healthcare decision making. *Proceed. CHI Confer. Human Factors Computing Syst.* 439, 1–17. doi: 10.1145/3613904.3642883

Kim, P. H., Dirks, K. T., and Cooper, C. D. (2009). The repair of trust: a dynamic bilateral perspective and multilevel conceptualization. *Acad. Manag. Rev.* 34, 401–422. doi: 10.5465/amr.2009.40631887

Kim, P. H., Dirks, K. T., Cooper, C. D., and Ferrin, D. L. (2006). When more blame is better than less: the implications of internal vs. external attributions for the repair of trust after a competence- vs. integrity-based trust violation. *Organ. Behav. Hum. Decis. Process.* 99, 49–65. doi: 10.1016/j.obhdp.2005.07.002

Kim, P. H., Ferrin, D. L., Cooper, C. D., and Dirks, K. T. (2004). Removing the shadow of suspicion: the effects of apology versus denial for repairing competence- versus integrity-based trust violations. *J. Appl. Psychol.* 89, 104–118. doi: 10.1037/0021-9010.89.1.104

Kim, T., and Song, H. (2021). How should intelligent agents apologize to restore trust? Interaction effects between anthropomorphism and apology attribution on trust repair. *Telematics Inform.* 61:101595. doi: 10.1016/j.tele.2021.101595

Langford, C. P. H., Bowsher, J., Maloney, J. P., and Lillis, P. P. (1997). Social support: a conceptual analysis. *J. Adv. Nurs.* 25, 95–100.

Lewicki, R. J., and Brinsfield, C. (2017). Trust repair. *Annual Rev. Organiz. Psychol. Organizational Behav.* 4, 287–313. doi: 10.1146/annurev-orgpsych-032516-113147

Li, Q., Luximon, Y., and Zhang, J. (2023). The influence of anthropomorphic cues on patients' perceived anthropomorphism, social presence, trust building, and acceptance of health care conversational agents: Within-subject web-based experiment. *J. Med. Internet Res.* 25, e44479. doi: 10.2196/44479

Li, S., Mou, Y., and Xu, J. (2025). Disclosing personal health information to emotional human doctors or unemotional AI doctors? Experimental evidence based on privacy calculus theory. *Int. J. Human–Computer Interaction* 41, 8593–8605. doi: 10.1080/10447318.2024.2411619

Liu, X., Xu, Z., Yu, X., and Oda, T. (2022). Why should I consult? The impact of social support on patient consultation in online healthcare communities. *Front. Psychol.* 13:993088. doi: 10.3389/fpsyg.2022.993088

Li, W., and Liu, X. (2025). Anxiety about artificial intelligence from patient and doctor-physician. *Patient Educ. Couns.* 133:108619. doi: 10.1016/j.pec.2024.108619

Li, X., Hess, T. J., and Valacich, J. S. (2008). Why do we trust new technology? A study of initial trust formation with organizational information systems. *J. Strateg. Inf. Syst.* 17, 39–71. doi: 10.1016/j.jsis.2008.01.001

Li, Y., Chen, L., and Fu, L. (2025). Vicarious interaction in online health consultation service: the effects of generative AI's anthropomorphism and social support on intended responses through social presence and source credibility. *Int. J. Human–Computer Interaction*. 41, 11209–11226. doi: 10.1080/10447318.2024.2441422

Madjar, N. (2008). Emotional and informational support from different sources and employee creativity. *J. Occup. Organ. Psychol.* 81, 83–100. doi: 10.1348/096317907X202464

Meng, H., Xiao, Q., and Na, Y. (2025). Warmhearted cues: a study of the impact of social mindfulness on trust repair by intelligent customer service in service recovery. *Int. J. Hosp. Manag.* 128:104131. doi: 10.1016/j.ijhm.2025.104131

Meng, J., and Dai, Y. (2021). Emotional support from AI chatbots: should a supportive partner self-disclose or not? *J. Comput.-Mediat. Commun.* 26, 207–222. doi: 10.1093/jcmc/zmab005

Nass, C., Steuer, J., and Tauber, E. R. (1994). Computers are social actors. *Proceed. SIGCHI Confer. Human Factors Computing Syst.*, 72–78. doi: 10.1145/191666.191703

Ohbuchi, K., Kameda, M., and Agarie, N. (1989). Apology as aggression control: its role in mediating appraisal of and response to harm. *J. Pers. Soc. Psychol.* 56, 219–227.

Oliver, R. L. (1980). A cognitive model of the antecedents and consequences of satisfaction decisions. *J. Mark. Res.* 17, 460–469.

Qin, X., Zhou, X., Chen, C., Wu, D., Zhou, H., Dong, X., et al. (2025). AI aversion or appreciation? A capability–personalization framework and a meta-analytic review. *Psychol. Bull.* 151, 580–599. doi: 10.1037/bul0000477

Quinn, T. P., Senadeera, M., Jacobs, S., Coghlan, S., and Le, V. (2021). Trust and medical AI: the challenges we face and the expertise needed to overcome them. *J. Am. Med. Inform. Assoc.* 28, 890–894. doi: 10.1093/jamia/ocaa268

Rashidi, F., Ghahremani, F., Mahmoodi, Z., and Doulabi, M. A. (2025). The role of social determinants of health in woman's intention to pregnancy: a model with the mediation of social support. *BMC Public Health* 25:1062. doi: 10.1186/s12889-025-22223-3

Reblin, M., and Uchino, B. N. (2008). Social and emotional support and its implication for health. *Curr. Opin. Psychiatry* 21, 201–205. doi: 10.1097/YCO.0b013e3282f3ad89

Ren, H., and Gray, B. (2009). Repairing relationship conflict: how violation types and culture influence the effectiveness of restoration rituals. *Acad. Manag. Rev.* 34, 105–126. doi: 10.5465/amr.2009.35713307

Schlenker, B. R., Pontari, B. A., and Christopher, A. N. (2001). Excuses and character: personal and social implications of excuses. *Personal. Soc. Psychol. Rev.* 5, 15–32. doi: 10.1207/S15327957PSPR0501_2

Sharma, K., Schoorman, F. D., and Ballinger, G. A. (2023). How can it be made right again? A review of trust repair research. *Aust. J. Manag.* 49, 363–399. doi: 10.1177/01492063221089897

Sitkin, S. B., and Roth, N. L. (1993). Explaining the limited effectiveness of legalistic "remedies" for trust/distrust. *Organ. Sci.* 4, 367–392.

Spreng, R. A., Harrell, G. D., and Mackoy, R. D. (1995). Service recovery: impact on satisfaction and intentions. *J. Serv. Mark.* 9, 15–23.

Sullivan, T. J. (1975). Achievement motivation and attribution theory. *Contemp. Sociol.* 4:425.

Sundar, S. S. (2008). *The MAIN model: A heuristic approach to understanding technology effects on credibility*. Cambridge, MA: MacArthur Foundation Digital Media and Learning Initiative, 73–100.

Tomlinson, E. C., Dineen, B. R., and Lewicki, R. J. (2004). The road to reconciliation: antecedents of victim willingness to reconcile following a broken promise. *Aust. J. Manag.* 30, 165–187. doi: 10.1016/j.jm.2003.01.003

Tomlinson, E. C., and Mayer, R. C. (2009). The role of causal attribution dimensions in trust repair. *Acad. Manag. Rev.* 34, 85–104. doi: 10.5465/amr.2009.35713291

Troshani, I., Rao Hill, S., Sherman, C., and Arthur, D. (2021). Do we trust in AI? Role of anthropomorphism and intelligence. *J. Comput. Inf. Syst.* 61, 481–491. doi: 10.1080/08874417.2020.1788473

Weiner, B. (1985). An attributional theory of achievement motivation and emotion. *Psychol. Rev.* 92, 548–573.

Wu, W., Deng, R., and Bai, L. (2025). How to apologize more effectively: the impact of trust repair timing and apology attribution on trust repair. *Curr. Psychol.* 44, 8865–8879. doi: 10.1007/s12144-025-07838-4

Zhang, X., Lee, S. K., Kim, W., and Hahn, S. (2023). "Sorry, it was my fault": repairing trust in human-robot interactions. *Int. J. Hum.-Comput. Stud.* 175:103031. doi: 10.1016/j.ijhcs.2023.103031

Zhou, C., and Chang, Q. (2024). Informational or emotional? Exploring the relative effects of chatbots' self-recovery strategies on consumer satisfaction. *J. Retail. Consum. Serv.* 78:103779. doi: 10.1016/j.jretconser.2024.103779