



OPEN ACCESS

EDITED BY

Ayodeji Olalekan Salau,
Afe Babalola University, Nigeria

REVIEWED BY

Antoine Khati,
University of Connecticut, United States
Raqiya Alharrasi,
Expert Verlag GmbH, Germany

*CORRESPONDENCE

Viana Nijia Zhang
✉ Nijia.Zhang2@newcastle.ac.uk

RECEIVED 27 June 2025
REVISED 14 October 2025
ACCEPTED 20 January 2026
PUBLISHED 04 March 2026

CITATION

Zhang VN, Vlachokyriakos V,
Johnson I and Durrant A (2026)
Understanding and responding to
complex online harms: misinformation,
fake news, and young adults.
Front. Comput. Sci. 8:1655186.
doi: 10.3389/fcomp.2026.1655186

COPYRIGHT

© 2026 Zhang, Vlachokyriakos, Johnson
and Durrant. This is an open-access
article distributed under the terms of the
[Creative Commons Attribution License
\(CC BY\)](#). The use, distribution or
reproduction in other forums is
permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication
in this journal is cited, in accordance
with accepted academic practice. No
use, distribution or reproduction is
permitted which does not comply with
these terms.

Understanding and responding to complex online harms: misinformation, fake news, and young adults

Viana Nijia Zhang*, Vasilis Vlachokyriakos, Ian Johnson and Abigail Durrant

Open Lab, School of Computing, Newcastle University, Newcastle upon Tyne, United Kingdom

Introduction: This United Kingdom (UK)-based study examines how online tools and technologies shape young adults' interactions with misinformation and fake news in everyday contexts, integrating insights from young adults and key stakeholders from both public and private sectors.

Methods: Through two data collection workshops—a stakeholder engagement session (N=22) and a co-design workshop with young adults aged 18 to 25 (M=7), we explored the challenges that young people face when encountering and interacting with misinformation and fake news online. Additionally, we examined the design of privacy-enhancing technologies, as well as the innovation and policy development priorities highlighted by our stakeholders.

Results: Our findings point to how young adults become vulnerable to exploitation by malicious actors online in various contexts, especially focusing on emotionally vulnerable life events. Our findings also emphasise the need for more empirical research that engages young adults within enclosed online communities, such as online gaming voice channels, where opinions can become radicalised, emotions intensified, and young adults desensitised.

Discussion: We propose implications for designing harm-reducing tools through increasing young people's individual agency, equipping them with the skills to recognise, assess, and address misinformation whilst also enhancing their algorithmic and new media literacy. We also advocate for increased reciprocal interactions and collaboration between mainstream and marginalised communities. These recommendations aim to guide the education sector, parents, policy-makers, media professionals, technology designers, and other stakeholders in exercising collective agency and fostering collaborative efforts to share communications and values that contribute to safeguarding a safer online environment for young adults.

KEYWORDS

misinformation, fake news, young adults, online harms, social media

1 Introduction

The social media landscape encompasses diverse web-based communication and networking tools (Wang et al., 2019). It emphasises user-generated content shaped by distinct social-cultural characteristics, influencing communication styles, values and ideologies worldwide (Butler, 2012; Chen et al., 2011; Ulrich, 1986). However, it has grown increasingly complex due to unregulated information flow and the spread of misinformation and fake news (e.g., Eysenbach, 2020; Freelon

and Wells, 2020). Social media and online communication platforms have become a fact of life for young adults (Young et al., 2023) in the United States (US) and globally. Their relationship with online misinformation and fake news is multifaceted, driven by their extensive platform use. In China, young adults can be not only targets but also spreaders, creators, or critics of false information (Chen et al., 2011). In the UK, Ofcom (2024b) reports that this demographic frequently encounters misinformation. Transnationally, misinformation is an established term that refers to a wide range of ‘false, misleading, and deceptive’ information ‘disorders’, varying by intention and veracity (Ecker et al., 2022). UK young adults use an average of 8.7 communication platforms, increasing their exposure to potentially harmful content (Ofcom, 2024a). Ofcom’s (2024a) *Adults’ Media Use and Attitudes Report* highlights that, despite high confidence in digital literacy, many remain unable to handle situations such as scam emails, leaving them vulnerable to misinformation and fake news.

The Online Safety Act 2023 (OSA) (Online Safety Act 2023, 2023) is a landmark piece of legislation designed to protect children and adults online by imposing duties to online user-to-user services—such as video-sharing platforms, forums, chat and dating services—to conduct risk assessments (Milmo, 2024). Currently, the OSA arguably lacks a credible and comprehensive plan to address the harms caused by online misinformation and fake news, leaving the public vulnerable to pervasive risks associated with online harms (Arshad and Brasted, 2024; Full Fact, 2024). Such harms include exposure to harmful or extremist ideologies, manipulation through false narratives, and victimisation by scams which have been exacerbated by the spread of online falsehoods (Elgot and Courea, 2024). Our study reported herein emerged within this context, responding to the gaps in the OSA by focusing on a demographic that continues to require societal attention.

Significant research and media attention have centred on children (e.g., Hopwood, 2023; Lepper, 2024) and teenagers (e.g., Almeida et al., 2023), particularly concerning their exposure to fake news (BBC, 2024) and the psychological impacts of manipulated imagery, such as beauty filters (Booth, 2024). However, young adults also warrant attention, as they are among the most active users of social media and video platforms (Anderson and Jiang, 2018). Reuters Institute report (Newman et al., 2022; Newman et al., 2023) show 39–44% of 18- to 24-year-olds ($N = 2000$) used Tiktok as their primary source of news between 2022 and 2023. By 2024, 86% of this group worldwide ($N = 10,024$) identified platforms including Instagram, Tiktok and X as main news sources (Newman et al., 2024). A UK survey published by The Alan Turing Institute and The Ada Lovelace Institute found that 41% of 18- to 24 year-olds reported the highest exposure to harmful online content (Enoch et al., 2023).

Despite growing independence and digital proficiency (Schwartz et al., 2015; Simpson, 2018), young adults remain susceptible to online misinformation, fake news, and harmful and illegal content due to their transitional developmental stage (Treadway and Holloway, 2018). Whilst aware of misleading content, many lack adequate knowledge of fact-checking tools and struggle to evaluate credibility (De Vicente Domínguez et al., 2021). As society increasingly relies on digital platforms, it is imperative to examine young adults’ exposure to misinformation and fake news, whilst addressing the challenges they face in navigating a deceptive digitalised information environment that transcends national boundaries.

This paper explores the complex online harms posed by misinformation and fake news affecting young adults in the UK. We focused on their lived experiences, complemented by stakeholder insights. We examined how young adults interact with social media content and exercise agency in identifying, navigating, and mitigating such harms. Whilst some literature treats ‘agency’ and ‘autonomy’ synonymously

(Bennett et al., 2023), others distinguish ‘autonomy’ as social-transformational and agency as the individual capacity to influence one’s environment (Kalaja and Rouhotie-Lyhty, 2021). We adopted the Human-Computer Interaction (HCI) definition of ‘agency’ as the experience of initiating intentional actions (Lukoff et al., 2021). Building on this, we examined the potential of harm-reducing tools to support young adults in taking control of their digital interactions, identifying misinformation and fake news, preventing its spread, and co-creating safer spaces whilst preserving freedom of social expression.

Our study had two aims. First, through engagements with stakeholders and young adults, we contextualised scenarios in which young adults encounter misinformation online and examined the associated harms and mechanisms. Second, through discussions and co-design sessions, we explored implications for developing privacy enhancing technologies (PETs) that reduce harms and support young adults in exercising agency and mitigating risks. Specifically, we addressed the following research questions.

RQ1: In what contexts are young adults more susceptible to online harms when navigating information?

RQ2: What are the implications for designing harm-reducing technologies that support young adults in exercising agency to identify online misinformation and fake news, whilst mitigating the harms enabled by such content?

We adopted a qualitative approach to embrace in-depth, idiographic insights from diverse stakeholder groups introduced above. Two workshops held in May and August 2023 engaged young adults—as primary social media consumers—alongside stakeholders from sectors responsible for safeguarding online information integrity. These workshops illuminated critical contexts and scenarios where complex harms emerge, as well as the roles individuals play in creating and disseminating them. We also investigated the vulnerabilities increasing young adults’ susceptibility to these harms. Our empirical engagements produced concrete, storyboard-based depictions of how young adults encounter misinformation in mundane settings. We further highlight the emotional dimensions of these interactions, the indicators used to assess trustworthiness, the vulnerabilities that increase their susceptibility to harm, and participant-proposed mitigation strategies. These findings provide valuable insights for designing harm-reducing tools to help young adults identify and evaluate online misinformation and fake news effectively.

We structure this paper as follows. In Section 2 we review the literature on young adults’ consumption of online misinformation and differing perspectives on agency. Section 3 details the study methods, participant demographics, and the three storyboards co-created by young adults during the second workshop. In Section 4 we present findings derived from the round-table discussions across both workshops, organised thematically. In Section 5 we reflect on these themes and participant contexts, proposing design implications for PETs that support content identification and the effective exercise of agency by young adults.

2 Background

2.1 Young people and misinformation

Although the literature often uses umbrella terms such as ‘young people’ or ‘youth’—often encompassing ages 15 to 35 (e.g.,

Duan, 2023; Ford et al., 2024; Juvalta et al., 2023)—this study focuses specifically on ‘young adults’ aged 18–25 due to their distinct cognitive and emotional characteristics (Ecker et al., 2022). This age marks a transitional phase involving identity formation, abstract reasoning, emotional reasoning, emotional regulation, and critical thinking (Schwartz et al., 2015; Simpson, 2018). As part of *Generation Z*, they are considered digital migrants shaped by technologies and frequently characterised as open-minded and homogeneous (Dimock, 2019; Insider Intelligence, 2023; Hassoun et al., 2023; Parker and Igielnik, 2020). Upon gaining legal adulthood, they can access age-restricted content and be responsible for criminal behaviours (Office for National Statistics, 2019), yet they remain susceptible to online misinformation due to ongoing cognitive development (Enoch et al., 2023; Quintas-Froufe et al., 2024).

Young adulthood is also a sensitive developmental period marked by enhanced vulnerability to environmental stressors (Biddle et al., 1990; Bleidorn et al., 2014). According to the UK’s ‘State of the nation 2022’ report, vulnerable young adults are often targets of coercion and exploitation (Department for Education, 2023), yet they may not recognise abuse or report it, engaging in high-risk behaviours (Arora et al., 2015; Bailey, 2017). Young et al. (2023) found that US college students perceive online harassment as pervasive but exhibit limited engagement in addressing it, often normalising hatred and promoting unjust social comparisons.

Despite awareness of misinformation and fake news, young adults often underestimate its impact. Gurgun et al. (2024) conducted a survey among 247 UK Facebook users and found they are less likely than older adults to challenge misinformation on the platform. A self-reporting survey during the COVID-19 identified young adults (aged 18–35) frequently spread misinformation, often without verifying content (Balakrishnan, 2024). Contributing factors include lack of media literacy education, time constraints, limited interest, and overconfidence in their critical skills (De Vicente Domínguez et al., 2021; Livingstone, 2004; Livingstone et al., 2019; Manca et al., 2021; Xiao et al., 2021).

The critical role of social media platforms in young adults’ communications has also been frequently highlighted in studies with this demographic during COVID-19 (Diepeveen and Pinet, 2022; Ford et al., 2024, 2025; List et al., 2015). Despite heavy use of these platforms, young adults regarded them as some of the least reliable sources of information (Diepeveen and Pinet, 2022). In examining the effectiveness of vaccine messaging from Canadian government Instagram accounts targeting 18–29 year olds for the COVID-19 vaccine, Ford et al. (2024) highlight that these efforts failed to meet the needs of this demographic, particularly in building trust and addressing online misinformation.

However, more research is needed to clearly define the contexts in which young adults encounter online misinformation and fake news, and how they make sense of it, beyond just understanding their general perceptions and attitudes. This will help identify specific risks in concrete scenarios and inform the development of more targeted interventions.

2.2 Online misinformation and correction

Yu et al. (2023) conducted a systematic review of 64 empirical studies to understand strategies for reducing misinformation

sharing and promoting correction sharing in today’s decentralised informational landscape. Their findings highlight several social media platform features that contribute to the recirculation of misinformation. These include credibility indicators such as user and expert reputation ratings (Kim et al., 2019); virality metrics like the number of retweets (Lee and Oh, 2017), likes and replies (Kim, 2018); the influence of power users (Shao et al., 2018); profit-driven algorithms (Braun and Eklund, 2019); and the employment of bots (Shao et al., 2018). The sequence of exposure to rumours and rebuttals also affects message perception and acceptance (Pal et al., 2020). These findings align with earlier research by Bryanov and Vziatyshva (2021), which explained why individuals fall for fake news and the factors influencing belief in misinformation. For instance, message topics and bandwagon heuristics, represented by the numbers of likes, can significantly enhance the perceived credibility (Luo et al., 2022). Yu et al. (2023) also found that perceived message importance mediates misinformation and correction-sharing behaviours (Oh and Lee, 2019), aligning with prior work on the role of innate beliefs and ideological predispositions (Pennycook and Rand, 2020). However, more research is needed to explore how young adults assess online information in everyday contexts and what actions they take based on trust or mistrust.

Existing studies in the social sciences and HCI have examined practical approaches to help a broader demographic identify and prevent the spread of misinformation and fake news online (e.g., Hassoun et al., 2023; Nazari et al., 2022; Staender and Humprecht, 2023; Veeriah, 2021). Various measures for fact-checking have been compared, designed, and evaluated, including the use of traditional news media, fact-checking organisations, professional journalists, as well as emerging approaches such as Artificial Intelligence (AI) and crowdsourcing (e.g., Almeida et al., 2023; Haque et al., 2020; Li and Chang, 2023; Liu et al., 2023; Santos and Pereira, 2024; Shusas, 2024). Sehat et al. (2024) proposed a priority framework to assist fact-checkers in making strategic decisions to evaluate and address the harms caused by misinformation. Other measures include interventions by trusted individuals, such as family members, who challenge the misinformed person offline—a strategy highlighted as effective by Scott et al. (2023). Additionally, games have been developed to raise awareness about the tactics employed by misinformation creators (e.g., Humphrey, 2023; Karlova, 2018; Maertens et al., 2021; Roozenbeek and Van Der Linden, 2019). Warning labels designed to discourage users from engaging with misinformation or assist in correcting it, along with preventive measures integrated into messaging platforms to limit content sharing, have also been explored (e.g., Broniatowski et al., 2023; Nassetta and Gross, 2020). However, some interventions, such as setting warning flags, may only be effective for a short period (Broniatowski et al., 2023; Scott et al., 2023).

Whilst researchers have started to explore specific vulnerable groups, such as older adults and individuals who have previously shared low-quality news (e.g., Brashier, 2024; Moore and Hancock, 2022; Peng et al., 2024), there remains a lack of evidence addressing the unique implications for young adults. Teenagers are growing up in an era where information on social media saturates their lives, and influencers significantly shape their options (e.g., Peter and Muth, 2023). We consider our work both timely and critical in raising awareness of these experiences. Furthermore, we offer considerations for stakeholders in global societies who are involved in the health and growth of younger generations.

2.3 Perspectives on agency

We define ‘agency’ as the subjectivity of individuals in holding motivational attitudes and taking initiative aligned with those attitudes (Hainz et al., 2016). In our study, we focus on young adults’ proactive attitudes and their capability to act against the spread of misinformation and fake news they encounter online. We emphasise the importance of their ability to shape outcomes rather than merely exchange information, whilst also evaluating such behaviours with regards to their own beliefs and attitudes (Hainz et al., 2016; Lukoff et al., 2021). Bennett et al. (2023) categorised human agency into four broad aspects: (1) the level and directness of casual involvement or value alignment (e.g., Birk et al., 2016; Lazar et al., 2017; Mentis et al., 2019; Valencia et al., 2020); (2) the sense of agency in triggering events (e.g., Kasahara et al., 2019); (3) the timescale of agency manifests, from momentary to lifelong (e.g., Cornelio Martinez et al., 2017; Foley et al., 2019); and (4) individual independence versus social interdependence (e.g., Garg, 2022; Güldenpfennig et al., 2019). Despite these categorisations, researchers often fail to articulate or examine the coordination and trade-offs between these dimensions (Bennett et al., 2023). Jicol et al. (2021), for instance, explored emotion and agency in Virtual Reality settings, showing that agency moderates the emotion’s effect on presence.

In the social sciences, particularly in Social Media Studies, agency has been examined through diverse lenses. Increasing attention is directed towards agency involving machines and algorithms in mainstream media and journalism (e.g., Linden, 2017; Rydenfelt, 2022; Verwiebe et al., 2024). Researchers have argued to maintain human oversight and control by ensuring individuals are informed about and capable of monitoring the behaviours of automated algorithmic agents, which should not independently make decisions (Linden, 2017; Rydenfelt, 2022). Complementing this perspective, Clark et al. (2015) explored the potential for agency in generation of young people through exchange storytelling. Their work highlighted how digitally enabled narrative exchanges on digital and social media platforms can foster a deeper awareness of the complex contexts that young people connect with.

Existing research across different fields is establishing connections. For instance, Bennett et al.’s (2023) generation of the interdependence within social contexts aligns with Clark et al.’s (2015) proposition of encouraging such awareness in education to generate agency among young people. However, a notable gap remains in further identifying other dimensions of agency that contribute to positive outcomes, such as enhancing individuals’ sense of self and wellbeing (Bennett et al., 2023; Lukoff et al., 2021). We recognise an opportunity through our recent work to contribute new, empirically grounded implications for how to promote agency in ways that enable young adults to protect themselves whilst mitigating the harms and risks associated with navigating information online, especially when encountering online misinformation and experiencing forms of online harm.

3 Methods

Our qualitative approach was grounded in a phenomenological methodology and our objective to empirically understand and evidence a ‘rich picture’ of individual experiences and perspectives from diverse cross-sector and cross-generational groups within the UK population (Durrant et al., 2018). This guided our sampling strategy

to engage diverse groups of individuals, and methods; we planned first a cross-sector workshop engaging a mix of professional stakeholders across job sectors plus young adults who have also professionally engaged with the subject; and a second ‘co-design’ workshop with young adults who also hold professional expertise or training in design or HCI.

The first stakeholder workshop (W1) engaged 22 stakeholders recruited through the research team’s professional network (see Table 1). The workshop aimed to gain a multi-faceted understanding from people who live or work closely with young adults, including the contexts and encounters with misinformation and fake news, and the related impacts.

The following young adult co-design workshop involved seven young adults with expertise in technology design and social sciences, who were also recruited through the research team’s professional network. Participants brought a relatively comprehensive understanding of how misinformation is created and disseminated through social media, along with enhanced insights into design principles, the needs of their peers, as well as the technical and practical feasibility of proposed solutions and challenges.

3.1 Phase one: expert stakeholder engagement

The sample comprised: nine external stakeholders with interdisciplinary backgrounds across private sector, NHS, Higher Education, the Voluntary, Community and Social Enterprise (VCSE) sector, and the public sector; two young adult representatives; and 13 internal project researchers. We held the workshop in a hybrid format (i.e., three offline groups and one online group), with a meeting room organised into roundtables. In accordance with the informed consent and workshop protocol, we assigned participant IDs to each participant along with a brief description of their role and expertise (see Table 1).

We engaged in discussions in response to three prompts across three round-table sessions. Each group appointed one researcher participant, to act as a coordinator and responsible for taking notes. At the end of each session, one participant from each group was selected to summarise the discussion outcomes on behalf of the group members. Following each presentation, all participants were encouraged to share their reflections.

After this workshop, we transcribed the audio recordings and collected hand-written notes captured at each table. We employed a Deductive Reflexive Thematic Analysis (RTA) method (Boyatzis, 1998; Braun and Clarke, 2012, 2019; Roberts et al., 2019), to capture, structure and interpret the data in relation to our predefined research questions, drawn from participants’ rich discussions and reflections (Braun and Clarke, 2006; Naeem et al., 2023). Audio recordings were transcribed by a trusted third-party company.

The first author of this paper began by thoroughly reviewing the transcripts to become familiarised with the data. The transcripts were then imported into NVivo for coding. Based on the existing research questions, the first author developed a set of preliminary codes and applied them to relevant segments of data. These codes were then subsequently organised and connected, facilitating the identification of recurring themes and patterns that aligned with the study’s focus.

The initial themes and corresponding coded excerpts were exported into an Excel file for further review and refinement. In line with RTA, the authors collectively reviewed the clustered themes

TABLE 1 Stakeholder workshop (W1) participant information.

Participant ID (W1)	Role	Expertise
P1	(Former) Police Representative	Design and deliver violence reduction themed learning with and for children and young people
P2	Public Media Corporation Representative	Data and Artificial Intelligence (AI) in news production and reporting
P3	Digital Identity Company Representative	Policy development for fraud prevention and safeguarding
P4	Undergraduate Student Representative	Computer Science
P5	Postgraduate Student Representative	Human-Computer Interaction (HCI) and Interaction Design
P6	Local, Black-Led Women Organisation Manager	Supporting black and minoritised women and children who have been subject to domestic and sexual violence
P7	Public Media Corporation Representative	Researching on public values and data-driven technology
P8	NHS Regional Integrated Care Board Representative	Public, patient, and carer engagement
P9	Local Cyber Security Cluster representative	Information security, cyber essentials, and strategic technology planning
P10	National Children Charity Group representative	Children Protection, working with children and young people in care.
P11	Local Youth Development Charity representative	Local youth engagement and service delivery
P12	Computer Science Co-Investigator	Applied machine learning
P13	Computer Science Researcher	Natural Language Processing
P14	Law Co-Investigator	Law and emerging technologies
P15	Law Researcher	Interaction of law with emerging technologies
P16	Ethics, Responsibility & Sustainability Co-Investigator	Developing theories of value to address sustainability issues
P17	Corporate Digital Responsibility (CDR) Researcher	CDR and digital ethics
P18	Computer Science Co-Investigator	Human-AI Interaction enhancement and AI-powered interactive systems
P19	Computer Science Researcher	Cyber security and Human-AI Interaction
P20	Media and Cultural Studies Co-Investigator	Media technologies and cultures
P21	Media and Cultural Studies Researcher	Digital media cultures, conflicts, and politics
P22	Finance Researcher	Information security, operational security for financial firms and governments

during weekly meetings, fostering critical reflection and ensuring that the themes accurately represented the research questions.

3.2 Phase two: co-design with young adults

In the second workshop (W2), the participants were assigned with participant IDs W2P1-W2P7. Our co-design method appropriated *Ideation Card Design Method* (Golembewski and Selby, 2010; Hope and Mulhall, 2024) and *Storyboarding Techniques* (e.g., Greenberg et al., 2012). Our objective was to facilitate the participants to collaboratively create (co-create) specific user scenarios of interacting with online technologies that demonstrate and depict their experiences mediated by online media consumption. Our appropriation of the Ideation Card method (Golembewski and Selby, 2010) utilised visual storyboarding (e.g., akin to comic strips) to depict the scenarios where individual (fictional) characters may encounter (or have encountered) online misinformation and fake news, and where interactions with online technologies can take place between particular people; these scenarios could be fictional

but inspired by or grounded in participants' personal experiences or the experiences of their peers. The rationale for using this method was to focus discussion on user experiences with technology (ibid). The broader rationale for the co-creative setting and use of storyboarded, character-driven scenarios was to prompt dialogue and reflection by young adults on specific contexts of interaction, to raise considerations for technology designers (Durrant et al., 2018) (see Table 2).

3.2.1 Card-based design method

Participants co-created 'Instance Cards' in a bespoke Ideation Card deck. We provided four card 'suits' in this deck: (i) *Character*, (ii) *Technology*, (iii) *(Mis) Information and Fake News*, and (iv) *Action*. Content to represent instances of three suits of cards [i.e., (i, ii, iii)] were co-created with the participants, whilst content for the 'Action' cards was provided by the researchers to scaffold ideation and ensure focus on relevant topics. The following prompt questions were also provided in conjunction with the theming of each card suit, to steer

TABLE 2 Co-design workshop (W2) participant information.

Participant ID (W2)	Gender	Affiliation
P1	Female	PhD student in HCI
P2	Male	Master student in HCI
P3	Male	PhD student in HCI
P4	Male	PhD student in HCI
P5	Female	Master student in HCI
P6	Female	Master student in HCI
P7	Female	User Researcher with an education background in HCI

conversation and reflection on: diverse personas; technologies involved in potential harms; and types of information encountered online.

- 1) *Character*—Who are we involving in this scenario? What is the character's name, age, and profession? What are they interested in? What are their routine activities?

(e.g., a 23-year-old man, John, who is a software engineer and part-time gamer. He goes to work from 9am till 5pm and plays games in the evening and during the weekends)

- 2) *Technologies*—What devices and online platforms are used when assessing the (mis/dis) information? Is the device or platform intelligent (i.e., using Machine Learning or Artificial Intelligent technologies)? What functional features does the device have?

(e.g., a desktop computer with Discord installed. It is not intelligent. It has vocal communication channels)

- 3) *Misinformation and Fake News*—What content or topic is being assessed? Who creates it (i.e., BBC journalist, TikTok users)? What form of multi-media is included (i.e., photo, video)? What population does such information target? What is the category of this information (i.e., misinformation that may be misinterpreted, or fake news or advertisements?)

(e.g., NHS-verified news but may be misinterpreted on social media, with a target of general public; fake news with an intention to mislead with its title and contents)

Each group of participants collaboratively created one instance card for each of the three suits. In addition to including illustrations, they provided detailed descriptions of each suit card, elaborating on the details (as shown in the right picture of Figure 1).

3.2.2 Storyboarding techniques

Authors 1 and 4 conducted the workshop. We split the participants into three groups (two or three people in each group). Each group picked one Instance Card from each suit, to form a card deck,

and then co-created a storyboard by adopting the information on the deck.

During the storyboard creation process, we encouraged participants to reflect on how the characters made sense of their encounters with information, and how engaging with the information made them feel. We also provided prompts that were inspired and informed by the Stakeholder Workshop (1) discussions, such as: 'What are the benefits of sharing such information?' and 'What are the consequences and harms?' After the creation of the storyboard, each group presented their story. Connecting with each story, participants reflected on the harms caused to young adults in similar real-life contexts.

We audio-recorded presentations of the three storyboards by each group, along with the subsequent reflective round-table discussions. These recordings were pseudonymised by the first author and transcribed by a professional transcription service. Given the narrative richness of data, which included both the narration of the storyboards and participants' discussions and reflections on the scenarios, we used Inductive RTA specifically to make sense of the discussion and reflection sessions (Braun and Clarke, 2012). The first author first familiarised herself with the data and organised transcripts and visual data for association with one of three distinct scenarios that were story-boarded. Participants' narrations and discussion of a given storyboard were analysed with the corresponding illustrations. Three distinct scenarios were generated: 'Hair growth advertisement'; 'Employment seeking scam'; 'Expert chatbot for health information'. The original language uttered by the participants in their narration was preserved throughout initial data processing. The first author then hand-coded the transcripts line-by-line using Microsoft Word. The codes were then critically discussed with the other authors at weekly meeting sessions and the first author reflected further and refined the codes. These were subsequently synthesised into initial themes for further collective analysis and consolidation by the research team, for example, to ascertain the sufficiency of data supporting each individual theme.

3.3 Co-design workshop storyboards

We now report on the storyboard contents and illustrations co-created and narrated by our participants in the co-design workshop.

3.3.1 Hair growth advertisement on Instagram

Gavin reads about 'the Miracle Anti-Hair Loss Treatment' Product on Instagram. He is always a bit anxious and worried about his appearance, and checks the product posts. Lots of people like the post. He also shares the post with his brother. In Gavin's eyes, his brother knows everything. His brother responds, 'Yes, it looks good. Try it!'. So, Gavin orders the product, and 4 years later, he is still as bald as his brother. He is sad because he could have used the money at somewhere else to make himself feel good.

3.3.2 Employment seeking advertisement scam

Chloe wants to look for a better job with more flexible time on her own. She encounters an ideal job on TikTok, which redirects her to the job advertisement link on Google. She applies for the job and receives a job offer. As she is preparing to quit her current job, suddenly she sees a post that claims the company she just received the offer is a scam. She says, 'Oh my! That was close'. She shares her experience to online communities and more people

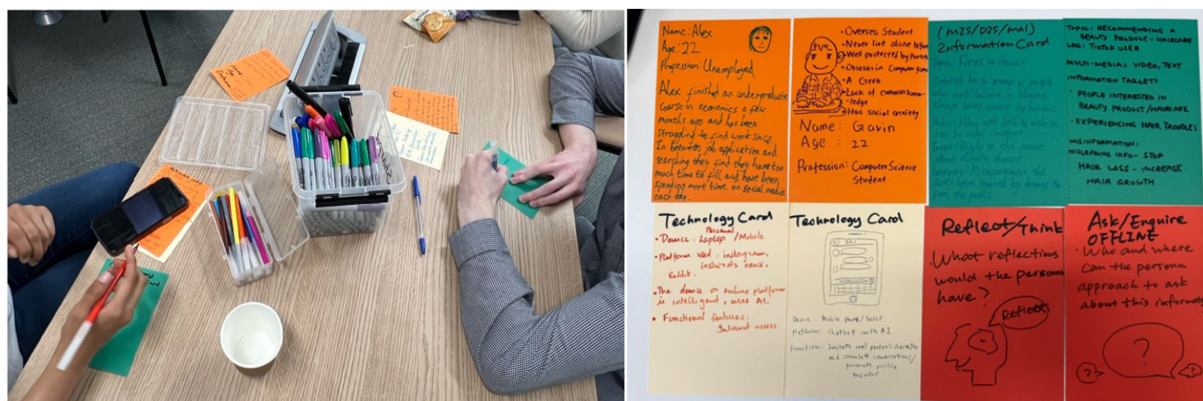


FIGURE 1

(Left) Participants each created 'information' instance cards. An example prompt card was placed on the table; (right) working in groups of two or three, participants chose one card from each 'suit' to collectively storyboard a scenario.

respond to her, mentioning the company is a scam to collect people's private data. Chloe does not know if she is going to be lucky next time.

3.3.3 An expert chatbot for health (mis) information

Steven is an undergraduate student majoring in Economics. Recently, he tested the Economy Chatbot built from ChatGPT. He started with small talks. The chatbot can understand him. He then tested the chatbot with economic concepts and to his surprise, the chatbot worked well. His trust on chatbot was increased. One day, when he is affected by a new form of virus—the Ninja virus, Steven goes to the chatbot for help. The chatbot is eager to figure out the symptoms and help him with potential treatment options...

Scenario One: Steven thinks it is still better for him to attend his GP service no matter what. He is given the right medicine and recovers in a few days.

Scenario Two: Steven decides to follow what the chatbot suggests and have a DIY treatment at home. The chatbot suggest vinegar having a powerful effect on this type of virus. Steven takes vinegar for a few days, but his situation gets much worse. He ends up in hospital (see Figure 2).

4 Findings

This following section presents the results of the deductive analysis of data recorded from the round-tables, and reflections from both workshops, addressing the research questions. These findings highlight how young adults encounter misinformation and fake news online and the related consequences. Additionally, the findings describe the individuals involved in the creation and dissemination of online misinformation and fake news to young adults and other younger generations, as well as the behaviours and strategies they employ in these processes. This has inspired our further translational work to deliver design implications for tools and resources that young adults should equip with in their future navigation online.

4.1 Emotional drivers in misinformation encounters

Our findings highlight that young adults' emotions and aspirations are often exploited in various contexts, plus their frustrations in seeking healthcare support, anxieties about fitting into social circles, efforts to secure employment, and desires for quick financial gains. In these situations, their eagerness to engage and succeed can increase their vulnerability to harm, particularly through the dissemination of misinformation by predatory 'bad actors'.

In discussing the spread of healthcare misinformation, W1P8 (NHS) described a scenario in which young adults 'get progressively frustrated by being passed from service provider to service provider', after 'following the traditional route to access a health service for a particular health problem'. The frustration and eagerness may drive them to seek out alternative support, which can result in serious harms. This NHS worker provided an example of 'people not getting a particular medication that they want prescribed, they may go and source it from alternatively online through illegal sources'.

We find that such urges to seek healthcare support can be taken advantage of by certain 'bad actors'. The public media corporation representative (W1P7) highlighted the allure of 'capitalism and profit-making', which draws in certain 'hidden game changers'—some pharmaceutical companies involved in collecting and monetising personal health data. These companies crafted targeted misinformation online, influencing and disrupting young adults' online efforts to seek reliable information and support.

Like the protagonist Gavin in the hair growth advertisement scenario, participants from the co-design workshop (W2) noted that young adults can be easily affected by social anxieties, and by the efforts made to fit in social circles. Hence, they tend to seek simple, 'One for all' (W2P3) solutions to their problems, frequently neglecting to invest time in scrutinising product descriptions or verifying the source of advertisements they encounter.

Social media platforms have arguably capitalised on the concerns and peer pressures experienced by young adults, by offering faster and more convenient pathways to online purchasing. As W2P4 noted, 'If you see it on a video or in a post, you can just literally click on a link, and it will take you to the shop that's within the social media app'. This can be exacerbated by 'celebrities and influencers' whose revenue depends on attracting audiences, promoting, and selling products (W1P12,

computer science). The selected publication channels, identified as 'social mediums', are often 'less inclined to moderate the content and advertisements'. W1P5 (postgraduate) added: 'Sometimes they are less inclined to moderate what advertisements they are actually selling through their promotions.' Arguably, such commercialised online environments further influence young adults' against applying logical judgment and critical thinking when navigating digital spaces.

In the 'Employment seeking advertisement scam' scenario, W2P5 highlighted that fake job scams would normally 'misplace letters' of renowned company names in order to target and mislead young adults who do not identify carefully when looking for a job, prompting them to think that the opportunity is with the renowned company. Refinement on the details such as reputation and history of fake companies would make it more difficult for young adults to distinguish



FIGURE 2 (Continued)



FIGURE 2
Storyboarded scenarios (top): hair growth advertisement on Instagram; (middle): employment seeking advertisement scam; (bottom): an expert chatbot for health (mis) information.

information veracity, because 'people thought it was reputable, so they didn't have any doubts' (W2P5). Young adults, in their desire to seize and secure job opportunities, may share personal information, such as identity and bank details, through the convenience of auto-complete browser functions. They seem to be susceptible to this when they want to secure a job position, and do not take the time to verify a company's legitimacy, potentially leading to financial loss, data breaches, and an increased risk of future 'scam calls'.

'When people are kind of desperate, I think people ignore all those kinds of signs like asking for money, things like that.' (W2P6)

W1P3 (Digital Identity Representative) cited the example of young adults being harmed by 'money muling'¹ advertisements, which can attract individuals with the promise of immediate financial gain whilst obscuring the potential long-term subsequent consequences for themselves and others:

'I was on a session with a young people's panel which was looking at the number of young people that just saw money muling as a way to make money, and they didn't understand the implications. They just say, "Oh, somebody's offering me £100, let me try this." They didn't understand what the risks were.'

1 <https://www.nationalcrimeagency.gov.uk/moneymuling> (Accessed January 2024).

Hence, W2P3 and W2P4 proposed that designers should consider how to stop young adults from making rushed decisions when what they 'normally look for efficiency'. Technology applications should be designed to prompt their users to make time to 'pause', to reflect on their own judgment, 'rather than being persuaded' (into a rushed decision).

W2P4 also observed that, everyone is susceptible to 'emotional moments' and may experience 'certain degree of disabilities' at times. Young adults in particular find themselves experiencing heightened emotional states, 'such as feeling isolated at home or about stressed job security and happiness' (W2P4). In such states, individuals are more likely to be harmed by misinformation and fake news, as they will 'feel difficult to identify them' (W2P4, W2P7). Consequently, they may be easily targeted or manipulated, as misleading information can offer 'a false sense of control' (W2P7). The emergence of AI algorithms raises new concerns, particularly around how new media technologies like chatbots may exploit young adults. These AI-driven technologies can project positivity, persuasiveness, and a 'keenness to help' (W2P1) by actively listening and providing support during young adults' expressed emotional needs. W2P3 commented that unlike 'a friend or a family member who may really want to help but don't have the information', such technologies can easily 'draw people in' with tailored responses.

In summary, this section reported how young adults' expressed emotions such as frustration, eagerness and anxiety can be exploited by targeted misinformation, especially during times of mental stress, physical discomfort, peer pressure, or financial instability.

Participants noted that such potential harm from misinformation encounters could extend to anyone experiencing heightened emotional states at certain points in life. In these contexts, misinformation targeting is shaped not only by content but also by underlying social dynamics.

4.2 Trust indicators of young adults

Across both workshops, participants agreed that young adults' trust in online information is highly susceptible to influence. Several contexts were described to illustrate the factors affecting trust, such as the perceived reliability of information, the influence of endorsements from trusted family members, the reactions of the public, and long-term influences by closed online communities dominated by singular and uniform opinions or voices.

In the co-design workshop (W2), young adults shared their curiosity about how peers assess the trustworthiness of health information. They observed that chatbot-generated responses might more readily earn users' trust because of the chatbot's anthropomorphic language use, making them more approachable and reliable. When searching for similar questions on search engine, participants noted that the results are often presented in an exaggerated manner. The outcomes can be fear-inducing and overwhelming.

"I think it's the ChatGPT virtual system, they made it sound human, so you actually trust it. Well people, if you search your condition on Google or on Bing it's just like, "Yeah, you're dead in like three seconds." (W2P2)

They also highlighted the difficulty of evaluating information provided by the ChatGPT chatbot. In particular, the 'one-time' (W2P2) and real-time nature of the communication makes it challenging for users to revisit and re-evaluate the previous responses provided as 'there's no coming back to questioning it' (W2P4). Hence, it is also difficult to lose the trust of human users once it proves its efficiency.

Trusted close family members, such as the brother in the 'hair growth advertisement on Instagram' scenario, can easily influence young adults' decisions to trust when faced with misleading advertisements, 'even though the (protagonist's) brother clearly doesn't have any knowledge around hair growth' (W2P3). To young adults, such trusted people can 'act as a proxy who is able to, hopefully, understand the world better than them', and let the demographic 'put down their guard'.

Young adults' trust to official information sources can be easily swayed, particularly when mistrust in health experts spreads rapidly through social 'contagion' (W1P14, law). For example, W1P14 spoke about 'vaccine hesitancy' during the Covid-19 pandemic as an example, describing how circulating doubts and mistrust towards health authorities led to widespread reluctance or refusal to get vaccinated. Additionally, public signals reflected by 'likes' or 'votes' on social media platforms (e.g., Instagram) can influence young adults' trust in online content. As W2P5 commented, 'some young people see the number of likes as a sort of, vote of confidence'.

Another high-risk context for young adults' exposure to online misinformation is within echo chambers. In these spaces, young adults' trust in and judgment of others can become desensitised, often leading to the normalisation of inappropriate behaviours. W1P12 (computer science) noted that individuals in echo chambers 'only got a partial awareness' and 'tend to only look at news or only hear the

views of people that agree with them'. Some young adults who 'believe everything is fine' will gradually stop questioning information validity.

In these closed communities, AI-generated 'personalised fake accounts' driven by 'political or societal campaigns' can effectively propagate misinformation to align with campaign goals. A significant challenge arises in tracking and preserving evidence within these real-time communication channels, making it challenging to counter misinformation. For example, W1P14 speaks about such an occurrence in the context of online gaming communities:

"There can be a lot of safeguarding issues around it, in part because if you're looking particularly at online gaming voice chat, you're talking about a lack of permanence, so it's often very difficult to have an evidence trail you can use to really identify who is at risk, who's being impacted, and how exactly this is being spread and disseminated." (W1P14)

Young adults' criteria for trust are found to be varied and often unpredictable. They may endorse trust in AI tools, close family members without specific expertise, or be swayed by public opinion, sometimes even challenging authorised and professional insights. Echo chambers can desensitise young adults to differing perspectives, making it harder to identify misinformation and engage with diverse viewpoints.

4.3 Vulnerabilities and susceptibility to information among young adults

Participants in the stakeholder engagement workshop highlighted several marginalised groups that are often overlooked by mainstream society and are particularly vulnerable to online misinformation. For instance, W1P14 (law) mentioned 'online gaming or virtual reality gaming' communities often marginalise or trivialise certain groups, particularly 'women or girls, and individuals from diverse racial backgrounds'. Such communities may perpetuate stereotypes or dismiss the experiences of these groups, reinforcing the exclusion of their voices. Hence, participants called for more inclusive research engagements with young adults whose voices are silent or silenced within online communities, to illustrate both idiographic journeys and experiences of complex online harms:

"It's remembering that there are populations of young people who are silent or silenced. We haven't come up with a solution as to how we reach those young people, but it's a case of there's something to think about." (W1P14)

W1P1 (Police) expanded the discussion of marginalised audiences to include not only young adults but also children. He highlighted concerns around 'an organised criminality harvesting the information of children, manipulating and engaging with them'. He raised an example of Kinder eggs to illustrate these tactics:

"Posting a photograph of Kinder egg with drugs² on an online platform. You then get children responding to that. So, the criminal has now harvested the contact details. He has identified

² For more details: <https://www.bbc.co.uk/news/uk-england-norfolk-41268211> (Accessed January 2024).

children who are interested in finding a Kinder egg full of drugs.’ (W1P1)

By identifying contact details of children who show interest through specific keyword searches on search engines, organised networks can effectively build a customer base. This enables stakeholders to ‘target their campaigns to children’ (W1P3) with particular vulnerabilities, minimising the need for direct targeting efforts.

In addition to specific gender, age, and racial groups, W1P3 (Digital Identity) highlighted the cascading chain effects of familial and social challenges faced by some young adults, which heighten their susceptibility to problematic situations. She noted that interconnected incidents are like ‘chinks in the armour’ that can exacerbate the fragility to misinformation and other online risks, such as being ‘*neurodiverse*’, experiencing ‘a parental break up’, or having ‘a parent or sibling going into prison’. Personal and social struggles often compound and limit their resilience and access to reliable support. She explained the reason:

‘Because if somebody has got a fragility, they’re the ones that are most likely to be targeted for different ends, be that through loneliness, be it through them having some gender or other element which singles them out.’

Malicious actors can exploit marginalised demographics with specific vulnerabilities by using AI algorithms to ‘ascertain characteristics such as gender, age, cognitive or physical disabilities’ (W1P3). This enables them to target misinformation more precisely at the selected audiences, increasing impact of misleading or harmful content.

Another form of vulnerability arises from the challenges of understanding complex medical, legal, or commercial terminologies online. Participants (W1P8 NHS, W1P14 law, W1P20 media) in the stakeholder engagement workshop (W1) emphasised that such language barriers prevent young adults from fully ‘understanding a situation’ and often leaving them ‘trying to make sense of in a given moment or situation’. They highlighted the example of young adults seeking healthcare support online. Given that online spaces provide healthcare information from diverse sources worldwide, often in different languages and cultural frameworks, the information exists in a legally ambiguous juridical territory. The space ‘is no longer governed by the best advice of a UK health provider’ (W1P14), making it challenging to safeguard young adults from unverified or potentially misleading healthcare sources, which may impact their wellbeing:

‘You may find advice that may be appropriate in the United States, where they have a different regulatory system or other forms of English language speaking.’

Young adults are vulnerable to spread of misinformation when online bystanders amplify the situation, particularly outside school spaces. W1P1 (Police) expressed concern about how content regarding conflicts between young adults can escalate and spread rapidly online. Online bystanders, such as ‘parents or just people who are on the periphery’, or those ‘just out of meanness who what to stir the pot’, can ‘add fuel to’ the fire, exacerbating the situation without considering the consequences. As W1P3 (Digital Identity) noted, people may not realise that what they are sharing is ‘fuelling, supporting and committing harmful conduct’.

In sum: across both workshops, participants identified several types of vulnerabilities that increase young adults’ susceptibility to misinformation. These include marginalised demographics, challenges in understanding complex legal, medical, or commercial terms across online boarders, and the potential harm caused by online bystanders outside of school-supervised environments. Such factors contribute to the potential risks young adults face online, as each can increase their exposure to and impact from misinformation.

4.4 Strategies for mitigating harms

To address the prevalence of misinformation, both W1P7 (Public Media) and W1P3 (Digital Identity) stressed the importance of equipping young people with personalised techniques and countermeasures to effectively ‘debunk and identify’ the harmful practices and strategies employed by bad actors. They highlighted the necessity to incorporate these resources into school curricula. Promoting media literacy, they argue, offers ‘a better way of trying to understand the spread of misinformation and who is impacted’.

W1P14 and W1P20 (Law and Media) highlight the importance of considering the ‘rationality and contextual nature’—the contextual significance of ‘agency’. They argue that solely increasing young adults’ agency is insufficient, as it can wrongly validate individuals’ perceived expertise on complex issues in which they personally lack knowledge, thus heightening their susceptibility to the influence of misinformation. Hence, the extent of agency granted to young adults in navigating online information and using digital devices requires reconsideration within societal disclosure:

‘We talked a little bit about that, and what does it actually mean to increase or decrease agency, and the relationality and contextual nature of agency itself basically, which makes it complicated to design for because it doesn’t necessarily mean that if you increase agency, that’s a good thing, because arguably, increasing agency was one of the reasons that misinformation was created in the first place.’ (W1P20)

Nevertheless, participants discussed how encouraging individuals to develop self-responsibility by promoting their agency is not the solution to mitigate against harms. In addition to endorsing individual agency, W1P20 (Media) also emphasised the importance of ‘collective agency’ (Hainz et al., 2016), which can foster group-based critical thinking about interventions and responses that can enable both informal and formal social controls. One example is the inclusion of oversight by mainstream journalists:

‘I think that collective agency (is a good point), because we’re always focusing on the individual. But this is where remit and responsibility through journalists’ work at the BBC, where you’re thinking about that all the time, the individual within the group.’ (W1P20)

The discussion of traditional gatekeepers in media accuracy such as news reporters and journalists prompted participants to explore evolving definitions of ‘guardians’ in the context of social media. W1P5 (postgraduate) noted that some young adults see themselves as guardians of safe online spaces, taking on roles in ‘advocating online safety’ and ‘providing peer support’. They actively monitor the veracity of information and exercise agency over how it is shared.

This theme highlighted participants' insights on key directions social media platform developers and policymakers should consider reducing young adults' exposure to and harm from misinformation. Discussions emphasised the importance of empowering young adults with agency to identify and stop the spread of misinformation, whilst also acknowledging challenges that may require additional intervention from mainstream gatekeepers and other stakeholders. In addition, providing more tailored materials to enhance young adults' media literacy skills in identifying misinformation was seen as critical.

5 Discussion

We reported in our Findings section three everyday contexts in which young adults may encounter and be adversely impacted by misinformation online. These contexts were co-created by young adults and highlighted how the fast-paced new-media environment can amplify this population's vulnerability to online information. Subsequently, we presented a qualitative data analysis of workshop discussions that shed light on emotional influences as key driver of misinformation and fake news encounters of young adults in their daily lives. Additionally, we examined the indicators young adults use to determine trustworthiness, the vulnerabilities that increase their susceptibility to misinformation, and the proposed strategies for mitigating the emotional, physical, and financial harms delineated in these contexts.

In this Discussion section, we deepen our analysis in relation to critical themes raised by participants in the workshops. These include the dynamics of echo chambers, the interplay between collective and individual agency, the cultivation of new media and algorithmic literacy (Chen et al., 2011; Koltay, 2011; Lin et al., 2013), strategies for future engagement with excluded voices, the design of slow thinking and emotion-buffering spaces, and approaches to support the emotional vulnerabilities of individuals in everyday settings. These insights offer valuable contributions to the fields both Social Sciences and Human-Computer Interaction design.

5.1 The 'malicious' echo chambers

During the first workshop (W1), stakeholders (e.g., W1P12) discussed the serious harm and risks posed by the spread of misinformation within small, enclosed online communities—often referred to as echo chambers or filter bubbles (e.g., Pennycook and Rand, 2020; Kaluža, 2022). Both terms are normally conflated in existing work and are frequently driven by social or political campaigns and exacerbated through the curation and manipulation of algorithms. Social media and other new media platforms afford both the representation and provocation of emotions, which can significantly influence the escalation and de-escalation of violence (Morales, 2024). The amplification of emotional responses and the resulting inappropriate behaviours, either through heated discussions or fabricated contents tailored specifically for the echo chamber spaces, can fuel conflicts by reinforcing polarised narratives.

An abundance of literature has explored and analysed how misinformation is created and disseminated to desensitise young adults within echo chambers. For instance, Diaz Ruiz and Nilsson (2023) highlighted that the malicious actors would strategically 'seed' the deceptions within these spaces by masquerading misinformation

as credible information. Once introduced, members within the echo chambers are enlisted to cocreate and propagate these contentious narratives, further disseminating the misinformation. As emphasised by W1P12, though initially seeking peers to share common interests, young adults may narrow their worldview to align solely with the perspectives within these spaces, leading to the subjective rejection of diverse voices and the gradual loss of critical thinking. However, there is a notable scarcity of empirical evidence involving users within echo chambers. Existing work, such as Söderberg's (2016) interviews with 13 users of anti-immigration internet media, explored why individuals consume such content and whether they are influenced by echo chambers. Despite this, it remains challenging to track and understand how young adults are gradually influenced within closed, encrypted groups like Telegram (Molla, 2021), where the dynamics of misinformation propagation and ideological reinforcement are often concealed. We propose adopting research methodologies such as data diaries (Vlachokyriakos et al., 2024), which allows for the longitudinal collection of self-reported data, to better track and understand young adults' behavioural change within such spaces over time. Notably, engaging with this demographic presents challenges, as young adults within echo chambers may be reluctant to participate in research activities or may resist acknowledging the presence of misinformation (Du, 2023, p. 20).

Participants also specifically talked about online gaming as a critical area for future research and engagement. The communication environment in gaming is often characterised by hostile and aggressive expressions, including discrimination, racism, and sexism (Meriläinen and Ruotsalainen, 2024). Due to the online disinhibition effect, young adults may feel less socially restrained, leading to more uninhibited and open expressions (Suler, 2004). Furthermore, the gaming environment can normalise such behaviours, making young adults perceive them as more acceptable (Meriläinen and Ruotsalainen, 2024). Compounding this issue, messages conveyed through voice are particularly difficult to capture as evidence, whilst tone and emotional delivery can more directly provoke responses and reactions (Lawson and Mayer, 2022). Given these dynamics, we believe this is an essential field requiring the involvement of researchers and educators to better understand and address the complexities within this unique context.

In addition, online technologies that enable encrypted communication and anonymity further complicate the situation within echo chambers by allowing individuals to engage in harmful activities without the risk of incrimination (Walther and McCoy, 2021). As W1P14 highlighted, AI-generated personalised fake accounts further contribute to the blend of voices in these enclosed spaces. Such AI deception—where AI systems are trained with malicious intent—can be seen as a process of human disempowerment (Park et al., 2024). In these echo chambers, intentionally created to attract vulnerable groups of young adults seeking peer support, voices gradually become contaminated by increasingly radicalised and polarised voices and opinions if they operate without supervision or moderation.

Hence, there is an urgent need to address how misinformation proliferates in these spaces and to evaluate the effectiveness of potential interventions through engagements with young adults and related stakeholders. By conducting empirical research in this area, researchers and related stakeholders can propose more effective policies and design tools that better safeguard young adults from these digital risks.

5.2 Implications

In this section, we consolidate insights from the workshops into strategic design implications.

5.2.1 Balancing collective and individual agency

Participants in the stakeholder-engagement workshop (W1) discussed both individual and collective agency. As highlighted by W1P20, solely increasing individual agency might inadvertently escalate the risks of introducing more misinformation. Furthermore, existing research suggests that enhancing agency in information search could also stem concerns about active intervention and customisations that may expose more personalised data (Rezk et al., 2024). Participants advocated for a critical perspective on collective agency. Drawing on Gonzalez Montero et al.'s (2024) arguments, we define collective agency as the integration of the casual powers of young adults and stakeholders, enabling the coordinated actions as a collective (Hainz et al., 2016). This encompasses their collaborative interactions with online information and the reciprocal effects and influences exerted on every receiver and reader of that information.

Building on our findings, we propose design implications for enhancing both individual and collective agency by incorporating several factors-informed agency. These align with the existing categorisations of agency perspectives, including emotion (Jicol et al., 2021), communication, and values, drawing on the foundation of existing work (e.g., Bennett et al., 2023; Hainz et al., 2016; Kavada, 2016; Rezk et al., 2024). Such perspectives can combine the interplay of individual and group efforts in navigating and mitigating misinformation and fake news online.

5.2.1.1 Emotion-informed individual agency

Crowd members are especially susceptible to blindly following others due to heightened emotions, which reinforces collective emotion states (Lebon, 2008). Participants in the co-design workshop (W2) highlighted that negative emotions among young adults are often exacerbated by the influence of celebrities and social media influencers. This emotional vulnerability, as highlighted in existing research, positions young adults for exploitation by malicious actors, political and social campaigns, as well as impulsive purchasing behaviours. Social media influencers, who convey social prestige, have emerged as new opinion leaders, experts, and brand ambassadors (Croes and Bartels, 2021; Lin et al., 2018). As discussed in Section 4.1, this individual fragility is further amplified within peer groups or echo chambers, where shared perspectives intensify emotional responses. Emotions such as anger can reinforce and perpetuate echo chamber dynamics, creating self-reinforcing cycles that deepen biases and limit young adults' exposure to diverse viewpoints (Wollebæk et al., 2019).

Young adults' emotions can be unconsciously triggered by bystanders—individuals who witness harmful actions or conflicts but either do not intervene or inadvertently escalate the situation (Mitchell et al., 2022). Participant W1P1 highlighted a widespread lack of public awareness regarding how sharing, twisting, or amplifying content can exacerbate conflicts among young adults. Research similarly emphasises the harmful role of online bystanders, whose actions can intensify tensions and extend real-world conflict into virtual realm, particularly on social media (Joelsson and Bruno, 2022). By unintentionally contributing to these dynamics, bystanders may heighten

emotional and psychological harm for young adults, both online and offline. This highlights the need for more regulatory action from social media platforms and legislative systems, as well as the urgent empowerment of young adults' individual agency in self-protection. Perpetrators and bystanders involved in bullying often become increasingly aware of the optimal time and places for harm, using spatial tactics to maximise social impact whilst minimising the possibility of 'being caught' (Andrews and Chen, 2006; Joelsson and Bruno, 2022). In response, young adults must develop critical awareness and resilience, employing self-protection strategies such as anonymous reporting and seeking support through accessible online and offline channels.

Institutional distrust is another emotional factor that can easily spread through online sources promoting pseudoscience and is closely linked to public anxieties (Han and Curtis, 2021). Mistrust of experts by certain individuals, exemplified by vaccine hesitancy as surfaced by participant W1P14 in our findings and reported substantially on similar work (e.g., Jaffe et al., 2022; Nurmi and Jaakola, 2023), can quickly spread to a broader audience. As participant W2P5 highlighted, mechanisms like public votes or likes on social media further amplify this mistrust, intensifying anxieties by moralising perceived 'risky' behaviours (Han and Curtis, 2021). To address the issue of 'collective anxiety' (Yang et al., 2021) and the spread of distrust, it is crucial to emphasise the role of individual agency. Existing research suggests that agency should only be prioritised when it can meaningfully contribute to the narrative and context of experiences (Jicol et al., 2021). Whilst this implication emerged in the context of VR settings, we argue that it is equally relevant to our study. Specifically, young adults' individual agency should be empowered and exercised when the awareness and ability to be independent, opt out of online communities or disengage from information sources provides them with a way to avoid being further influenced by collective emotions (Bennett et al., 2023). This is particularly important in situations where they feel overwhelmed or when such interactions negatively impact their emotional wellbeing.

5.2.1.2 Communication and value-informed collective agency

The adoption and cultivation of collective agency also requires the efforts and power from news publishers, curators, and journalists, as emphasised by participant W1P20, beyond the individualistic approaches (Aghajari et al., 2023). From the young adults' perspective, this is important because the processes of self-control and self-change can both support and constrain the endorsement of agency over time (Bennett et al., 2023; Lyngs et al., 2022). Young adults' goals, values, and intentions can vary, and as some younger children transition into young adulthood every day, the delegation of agency requires not only oversight by schools but also involvement from gatekeepers, such as publishers and journalists, whose voices are increasingly undermined by the risks of algorithms (Bhuiyan et al., 2021; McClure Haughey et al., 2020). Whilst from a societal perspective, we need collective agency as it reflects the values derived from collaborative communications and actions among all related stakeholders (Kavada, 2016). These stakeholders include not only traditional and new social media content curators but also parents and teachers, Internet influencers, and policymakers—individuals who must recognise the effectiveness and legitimacy of rules and principles aimed at preventing the spread of misinformation or fake news (Hainz et al., 2016; Mathiesen, 2006).

Online media can be regarded as one of the sites for conversations and textual information exchange, offering affordances that enable interactions among diverse stakeholders—facilitating the sharing and exchanging of opinions, as well as evaluations on the information consumed and disseminated (Kavada, 2016).

By acknowledging these standards and committing to act in accordance with them, these stakeholders collectively contribute to fostering a responsible and informed online environment. Such efforts can also promote societal collective beliefs and desires to safeguard a safe and responsible digital space that satisfies the needs of young adults who navigate information online (List and Pettit, 2011; Mathiesen, 2006). In many everyday situations, individuals act not in isolation but within social contexts (Moore and Obhi, 2012). We believe that the collective agency, in turn, reinforces and stimulates young adults' individual agency. This reciprocal dynamic aligns with the concept of 'interdependence' within social contexts—a perspective of agency highlighted in Bennett et al.'s (2023) work—where collaborative actions within the social frameworks strengthen individual capacities for informed actions.

5.2.2 Cultivating new media and algorithmic literacy

Professionals in the first workshop emphasised the growing concern about young adults' vulnerability to unintentionally or unknowingly entering legally binding contracts and illegal sources of healthcare online. With the increasing prevalence of digital devices, children and young adults face greater risks of exposure to legal obligations and unauthorised healthcare or mediations without full understanding the implications or professional languages (Blum-Ross and Livingstone, 2017; Wszalek, 2017). Hence, promoting education about the legislative frameworks and authorised healthcare practices can contribute to raising awareness and improving understanding among children and young adults. In accordance with the Online Safety Act 2023 (2023), both demographics should be equipped with 'sufficient capability' to evaluate online contents and services that they attend to. They should be empowered with the knowledge and tools to identify underlying harmful online interactions. In addition, they should also be educated about their responsibilities and rights to report the 'abusive contents', 'hatred' behaviours, and illegal bindings online.

We would also like to highlight an intriguing avenue for future research based on insights from participating young adults in the co-design workshop (W2). They described mundane scenarios in which ChatGPT becomes seamlessly integrated into their daily communications, information searches, and decision-making. The co-created storyboard illustrates a situation where young adults carefully test and assess ChatGPT's credibility before trusting it with more aspects of daily decision-making. Recent research examining scientific papers with signs of GPT use on Google Scholar, particularly those focused on health-related topics, found that 57% of the selected questionable papers were susceptible to influence operations (Haider et al., 2024). Such convincingly scientific looking paper would undermine the basis for trust in scientific knowledge, bringing detrimental harm to the trust of young adults in academic professionals and institutions.

In the era of social media, these 'mighty' AI algorithms have become significant factors in shaping how young adults perceive and process information. AI algorithms now act as a new type of 'actor' that intervenes in communication and influences individual lives

(Shin et al., 2022). They bring convenience to media consumers, delivering personalised information tailored to individual interests and preferences (Li et al., 2011). However, the lack of transparency and non-deterministic nature of some AI algorithms poses challenges in identifying the sources and pathways of information, as it may sometimes be fully fabricated to maximise relevance based on the audience's perceived characteristics (Du, 2023). Whilst algorithms may present certain facts accurately, their inherent opacity can foster mistrust, paradoxically leading to an over-reliance on human oversight and intervention in the technology. This phenomenon, referred to as the 'the automation paradox' in existing work, suggests that as automated systems becomes more efficient, the need for meaningful human involvement becomes increasingly critical (KPI, 2024). This dynamic poses potential risks to young adults, as excessive intervention or over-reliance could inadvertently exacerbate harms rather than mitigate them. In addition, existing literature has examined how algorithm-driven recommendations and push content may undermine young adults' sense of agency in selecting precepted information by continually reinforcing and delivering content lined to negative keywords (Lukoff et al., 2021).

Hence, we emphasise the importance of considering the role AI algorithms play in the social media landscape and the need to enhance individual's algorithmic literacy—the awareness, knowledge, imagined understanding, and tactics that are responsible for content personalisation mechanisms (Bell et al., 2023). This also includes promoting skills in both algorithm appreciation and avoidance (Joris et al., 2021; Shin, 2020). To assist young adults in detecting and reporting misinformation, credibility tools should be enhanced to address the challenges posed by 'deepfakes' and 'cheap fakes'³ (Gamage et al., 2022). Value-centric curatorial algorithms and social media systems should also be designed and developed that can embed values such as prioritising children and young adults' emotional wellbeing beyond engagement.

5.2.3 More engagement with excluded voices

Certain groups of 'strategic actors'—such as activists ranging from large unions to small citizen groups are overrepresented in the social media environment (Chalmers and Shotton, 2016; Humprecht et al., 2020). Attention should also shift towards amplifying silent voices and exploring less traditional digital interaction spaces, such as voice chat channels and emerging video game platforms as addressed in Section 4.1. These platforms offer unique environments where marginalised groups may communicate or express themselves in ways that are often overlooked in mainstream studies. There is also existing proposed solution to amplify these voices is the use of social bots, which can artificially inflate the number of likes and shares, creating a manufactured sense of popularity (Bradshaw and Howard, 2017). Whilst this may bring attention to marginalised ideas, it raises ethical concerns about authenticity and the potential manipulation of online disclosure (e.g., Braddock et al., 2024). We need to explore effective ways to closely engaged with underrepresented voices, whilst ensuring the efforts to elevate them do not inadvertently distort the integrity of the dissemination of online

³ 'Deepfake' refers to 'videos that use some form of deep or machine learning to hybridise or generate human bodies and faces'. 'Cheap fake', as another form of Audiovisual Manipulation, refers to 'videos created relying on cheap, accessible software, or no software at all' (Paris and Donovan, 2019).

communication. There are also proposals for future research to incorporate Community-based Participatory Design and collaborative design approaches (Fathallah, 2021; Harrington et al., 2019), or offering solutions such as establishing local libraries and encouraging storytelling within communities (Amazeen et al., 2024). These initiatives aim to foster civic and community engagement whilst empowering marginalised individuals throughout the process of involvement.

Workshop participants did not offer additional strategies for engaging with marginalised groups. However, we want to extend beyond existing approaches, highlighting that solely recruiting marginalised participants into the design and engagement process may be insufficient for addressing misinformation, as it primarily emphasises one-way inclusions. Instead, we believe there is a need for fostering meaningful, reciprocal communications, where voices are not only heard but actively engaged in efforts to achieve shared understanding. Recent research suggests that regular interactions with minoritised groups can help bridge the radical divides surrounding important current events (Abrajano et al., 2024). Active, interpersonal engagement with individuals from either similar or different minoritised backgrounds is positively associated with correcting misconceptions. Therefore, we propose that, beyond general strategies to amplify excluded voices, fostering mutual interactions between marginalised and mainstream voices especially among young adults could help individuals understand how misconceptions and misinformation form about or within both mainstream and marginalised communities.

5.2.4 Designing slow thinking and emotional buffering for young adults

The concept of designing a ‘slow thinking space’ was inspired by W2P3 and W2P4’s comments as a strategy to promote more thoughtful decision-making, particularly when individuals are at high risk of exposure to misinformation. This heuristic aims to help peers make more informed decisions and reduce the risks associated with harmful content. Additionally, it seeks to empower young users as they navigate online environments by encouraging thoughtful, deliberate engagement.

Fast and slow thinking modes can be distinguished and measured by the time spent on decision-making (Jimenez et al., 2018). Fast thinking is characterised by factors including speed, intuition, emotion responses, and low awareness; whilst slow thinking involves careful interpretation, analysis, heightened awareness, doubt, and stress (Kahneman, 2012). Recent research has shown that individuals aged 14 to 19 are primarily characterised by a fast-thinking style (Jabbar et al., 2024). Young adults are in a stage of development where slow thinking becomes more prominent, yet they remain vulnerable, especially when consuming social media. Research has shown that users tend to think more intuitively and make quicker intuitive decisions on a smartphone compared to a personal computer (PC), increasing their susceptibility to impulsive judgments.

Research by Baum and Abdel Rahman (2021) further underscores the complexity of these cognitive processes, revealing that simply reminding individuals of the trustworthiness of media sources before exposure to information does little to counteract the effects of emotionally charged headlines. Negative social-emotional information, even from untrustworthy sources, is participationally hard to ignore. Fast and slow cognitive responses to such headlines remain biased, as emotions often override critical thinking and reflection, regardless of whether individuals engage in fast or slow thinking.

In addition to designing a ‘slow thinking space’, we advocate for future research to explore the creation of an ‘emotion buffer space’. This concept would protect young adults from being immediately confronted with emotionally charged headlines, providing them with the cognitive agency to pause and engage in more critical judgment. By allowing more time for reflection, this approach could help prevent emotional responses from completely dominating their information evaluation and decision-making processes.

5.2.5 Support for emotionally vulnerable moments in life

In the first workshop, participant W1P8 emphasised how frustration can lead young adults to seek out ‘illegal sources of health providers’, potentially resulting in serious harm. This concern was echoed by the young adult designers in the second workshop, who highlighted that everyone can feel vulnerable and isolated at certain points in life, making it challenging to identify reliable information. Emotional vulnerability is closely tied to social discomfort and reflects an individual’s sensitivity and the likelihood of getting hurt (Khrystenko, 2022; Yamaguchi et al., 2022). People tend to focus more on information that worries them rather than information that does not (MacLeod et al., 1986). When they are highly sensitive and exposed to more concern-related information, they may be especially vulnerable to the harmful effects of mis- or dis-information, particularly when algorithms on social media push this type of content to them.

5.3 Limitations and future research

Our qualitative approach for both workshops determined a relatively small, non-representative UK study sample and could be complemented by a larger-scale future study with a representative sample. Also, in the co-design workshop (W2), our choice of co-creation methods led to our recruitment of participants who have specialist expertise in Interaction Design. This focus on expert participants introduced a bias, because it engaged young adults with higher levels of media literacy and technology skills. Future research could involve a more diverse sample of young adults.

In terms of extending the qualitative research on this topic and reflecting on methods: whilst the storyboarding techniques drew anecdotally on lived experiences, there is also an opportunity to use different speculative design methods that engage research populations in the real-world contexts of their lives, for example, more performative techniques (Elsden et al., 2017; Hutchinson et al., 2025), or other real-world interventions and observations (Cipolla, 2018). Such methods could invite co-creativity and speculation in participants, whilst also capturing direct experiences of media creation and consumption that could deliver evidence on situated behaviours and sense making.

6 Conclusion

This qualitative study explored broad and specific contexts in which young adults living in the UK become vulnerable to encountering and being affected by online misinformation and fake news. We adopted a qualitative, idiographic approach to engage diverse stakeholders including young adults in rich roundtable discussions, supported by workshops that utilised co-creative methods (including visual storyboarding) to evidence individual experiences and perspectives. We explored how

emotions may prompt and shape misinformation encounters, and how social factors influence young adults' trust in various online sources. We further illuminated the vulnerabilities that malicious actors exploit, plus participant-proposed mitigation strategies.

We address the challenges of engaging and supporting young adults in echo chambers and the importance of safeguarding them from additional harms within these enclosed online communities. Reflections on ideas of 'agency' are offered from both individual and collective perspectives, reflecting on how factors such as emotions, communication, and values shape not only young adults' individual sense of agency but also the collaborative agency of related stakeholders, who should uphold the veracity and accuracy of online information. Additionally, we stress the criticality of cultivating new media literacy, particularly algorithmic literacy, placing focus on how AI-driven chatbots and other emerging technologies shape young adults' information assessment and decision-making processes. We highlight vulnerabilities such as marginalised demographics and the challenges in understanding professional terminology. Recognising the emotional vulnerabilities inherent to everyone—not only young adults—underscores the need to protect online technology users against exploitation by malicious actors; nudges may be designed into technology applications to discourage rash and rushed decision making and to encourage critical reflection-in-action.

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 SAGE Ethics Committee of Newcastle 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. 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

VZ: Resources, Visualization, Conceptualization, Formal analysis, Project administration, Software, Data curation, Methodology, Writing – review & editing, Writing – original draft. VV: Methodology, Supervision, Validation, Writing – review & editing, Project administration, Formal analysis, Funding acquisition. IJ: Formal analysis, Validation, Writing – review & editing. AD: Resources, Funding acquisition, Project administration, Visualization, Data curation, Validation, Writing – review & editing, Supervision, Conceptualization, Investigation, Methodology.

Funding

The author(s) declared that financial support was received for this work and/or its publication. This research is supported by both the AGENCY Project and the Centre for Digital Citizens (CDC) Project. 'AGENCY: Assuring Citizen Agency in a World with Complex Online Harms' is an interdisciplinary research project (<https://agencyresearch.net/>) funded by EPSRC through REPHRAIN, the National Research Centre on Privacy, Harm Reduction and Adversarial Influence Online (<https://www.rephrain.ac.uk/>). The aim of the AGENCY project is to explore the design of tools that empower their users to manage the potential risk of online harms and address unmet user needs within the UK population. The project combines academic expertise in Computer Science, Design, Communication Studies, Business, Applied Ethics and Law and involves applied (real-world) research with stakeholders and intended end users of online technologies, and sketching and evaluation of novel design concepts to inform the development of privacy enhancing technologies. Furthermore, AGENCY is case study based, to address particular contexts of interaction in which complex harms may take place. CDC centre is also funded by EPSRC. It takes an inclusive, participatory approach and works with citizens and partners to co-design sustainable 'Digital Social Innovations' that ensure technologies applications support diverse communities and have long-lasting social value and impact beyond the life of the Centre (<https://digitalcitizens.uk>).

Conflict of interest

The author(s) declared that this work 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) declared that Generative AI was not 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

- Abrajano, M., Lajevardi, N., and Uribe, L. (2024). Context, contact, and misinformation about socially marginalized groups in the United States. *J. Race Ethn. Polit.* 9, 335–376. doi: 10.1017/rep.2023.44
- Aghajari, Z., Baumer, E. P. S., and DiFranzo, D. (2023). Reviewing interventions to address misinformation: the need to expand our vision beyond an individualistic focus. *Proc. ACM Hum.-Comput. Interact.* 7, 1–34. doi: 10.1145/3579520
- Almeida, C., Macedo-Rouet, M., de Carvalho, V. B., Castilhos, W., Ramalho, M., Amorim, L., et al. (2023). When does credibility matter? The assessment of information sources in teenagers navigation regimes. *J. Librariansh. Inf. Sci.* 55, 218–231. doi: 10.1177/09610006211064647
- Amazeen, M. A., Vasquez, R. A., Krishna, A., Ji, Y. G., Su, C. C., and Cummings, J. J. (2024). Missing voices: examining how misinformation-susceptible individuals from underrepresented communities engage, perceive, and combat science misinformation. *Sci. Commun.* 46, 3–35. doi: 10.1177/10755470231217536
- Anderson, M., and Jiang, J. (2018). Teens, social media & technology 2018 (No. 202.419.4372). Washington, DC: Pew Research Centre.
- Andrews, G. J., and Chen, S. (2006). The production of tyrannical space. *Child. Geogr.* 4, 239–250. doi: 10.1080/14733280600807120
- Arora, S., Shah, D., Chaturvedi, S., and Gupta, P. (2015). Defining and measuring vulnerability in young people. *Indian J. Community Med.* 40, 193–197. doi: 10.4103/0970-0218.158868
- Arshad, T., and Brasted, C. (2024). What's next for the UK's online safety act and can it solve the misinformation problem? Available online at: <https://www.hoganlovells.com/en/publications/whats-next-for-the-uks-online-safety-act-and-can-it-solve-the-misinformation-problem> (Accessed December 22, 2024).
- Bailey, S. (2017) Unite against child sexual exploitation [National Police Chief's Council]. Available online at: <https://news.npcc.police.uk/releases/cc-simon-bailey-blog-unite-against-child-sexual-exploitation> (Accessed December 22, 2024).
- Balakrishnan, V. (2024). Socio-demographic predictors for misinformation sharing and authenticating amidst the COVID-19 pandemic among Malaysian Young adults. *Inf. Dev.* 40, 319–331. doi: 10.1177/02666669221118922
- Baum, J., and Abdel Rahman, R. (2021). Emotional news affects social judgments independent of perceived media credibility. *Soc. Cogn. Affect. Neurosci.* 16, 280–291. doi: 10.1093/scan/nsaa164
- BBC. (2024). Half of teenagers worried about fake news. Available online at: <https://www.bbc.co.uk/bitesize/articles/zsrtp2p> (Accessed December 22, 2024).
- Bell, A. R., Tennfjord, M. K., Tokovska, M., and Eg, R. (2023). Exploring the role of social media literacy in adolescents' experiences with personalization: a Norwegian qualitative study. *J. Adolesc. Adult. Lit.* 66, 239–248. doi: 10.1002/jaal.1273
- Bennett, D., Metatla, O., Roudaut, A., and Mekler, E. D. (2023). How does HCI understand human agency and autonomy? *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, 1–18
- Bhuiyan, M. M., Whitley, H., Horning, M., Lee, S. W., and Mitra, T. (2021). Designing transparency cues in online news platforms to promote trust: journalists' & consumers' perspectives. *Proc. ACM Hum.-Comput. Interact.* 5, 1–31. doi: 10.1145/3479539
- Biddle, B. J., Bank, B. J., and Slavings, R. L. (1990). Modality of thought, campus experiences, and the development of values. *J. Educ. Psychol.* 82, 671–682. doi: 10.1037/0022-0663.82.4.671
- Birk, M. V., Atkins, C., Bowey, J. T., and Mandryk, R. L. (2016). Fostering intrinsic motivation through avatar identification in digital games. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, 2982–2995
- Bleidorn, W., Kandler, C., and Caspi, A. (2014). The behavioural genetics of personality development in adulthood—classic, contemporary, and future trends. *Eur. J. Personal.* 28, 244–255. doi: 10.1002/per.1957
- Blum-Ross, A., and Livingstone, S. (2017). “Sharenting,” parent blogging, and the boundaries of the digital self. *Popul. Commun.* 15, 110–125. doi: 10.1080/15405702.2016.1223300
- Booth, R. (2024). “Teenage girls are feeling vulnerable”: Fears grow over online beauty filters. The Guardian. Available online at: <https://www.theguardian.com/media/2024/nov/29/teenage-girls-are-feeling-vulnerable-fears-grow-over-online-beauty-filters> (Accessed December 22, 2024).
- Boyatzis, R. E. (1998). Transforming qualitative information: Thematic analysis and code development. Thousand Oaks, CA: Sage Publications.
- Braddock, K., Hughes, B., Goldberg, B., and Miller-Idriss, C. (2024). Engagement in subversive online activity predicts susceptibility to persuasion by far-right extremist propaganda. *New Media Soc.* 26, 1775–1798. doi: 10.1177/14614448221077286
- Bradshaw, S., and Howard, P. N. (2017) Troops, trolls and troublemakers: A global inventory of organized social media manipulation (computational propaganda research project) [Working paper], University of Oxford. Available online at: <https://ora.ox.ac.uk/objects/uuid:cef7e8d9-27bf-4ea5-9fd6-855209b3e1f6/files/m3ca8c455852611e82d0fb182445a471f> (Accessed December 22, 2024).
- Brashier, N. M. (2024). Fighting misinformation among the most vulnerable users. *Curr. Opin. Psychol.* 57:101813. doi: 10.1016/j.copsy.2024.101813
- Braun, V., and Clarke, V. (2006). Qualitative research in psychology using thematic analysis in psychology using thematic analysis in psychology. *Qual. Res. Psychol.* 3, 77–101. doi: 10.1191/1478088706qp063oa
- Braun, V., and Clarke, V. (2012). “Thematic analysis” in APA handbook of research methods in psychology, eds. Cooper, H., Camic, P. M., Long, D. L., Panter, A. T., Rindskopf, D., and Sher, K. J. Vol 2: Research designs. Washington, DC; Hershey, Pennsylvania; London: Quantitative, qualitative, neuropsychological, and biological, vol. 2, 57–71.
- Braun, V., and Clarke, V. (2019). Reflecting on reflexive thematic analysis. *Qual. Res. Sport Exerc. Health* 11, 589–597. doi: 10.1080/2159676X.2019.1628806
- Braun, J. A., and Eklund, J. L. (2019). Fake news, real money: ad tech platforms, profit-driven hoaxes, and the business of journalism. *Digit. Journal.* 7, 1–21. doi: 10.1080/21670811.2018.1556314
- Broniatowski, D. A., Simons, J. R., Gu, J., Jamison, A. M., and Abrams, L. C. (2023). The efficacy of Facebook's vaccine misinformation policies and architecture during the COVID-19 pandemic. *Sci. Adv.* 9:eadh2132. doi: 10.1126/sciadv.adh2132
- Bryanov, K., and Vziatyshva, V. (2021). Determinants of individuals' belief in fake news: a scoping review determinants of belief in fake news. *PLoS One* 16:e0253717. doi: 10.1371/journal.pone.0253717
- Butler, J. (2012). “Grappling with change: web 2.0 and teacher educators” in Developing technology-rich teacher education programs. Washington, DC; Hershey, Pennsylvania; London: Key issues, Vol. 1, 135–150.
- Chalmers, A. W., and Shotton, P. A. (2016). Changing the face of advocacy? Explaining interest organizations' use of social media strategies. *Polit. Commun.* 33, 374–391. doi: 10.1080/10584609.2015.1043477
- Chen, D.-T., Wu, J., and Wang, Y.-M. (2011). Unpacking new media literacy. *Syst. Cybern. Inform.* 9, 84–88.
- Cipolla, C. (2018). Designing for vulnerability: interpersonal relations and design. *She Ji: J. Design Eco. Inn* 4, 111–122. doi: 10.1016/j.sheji.2018.03.001
- Clark, W., Couldry, N., MacDonald, R., and Stephansen, H. C. (2015). Digital platforms and narrative exchange: hidden constraints, emerging agency. *New Media Soc.* 17, 919–938. doi: 10.1177/1461444813518579
- Cornelio Martinez, P. I., De Pirro, S., Vi, C. T., and Subramanian, S. (2017). Agency in mid-air interfaces. *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 2426–2439
- Croes, E., and Bartels, J. (2021). Young adults' motivations for following social influencers and their relationship to identification and buying behavior. *Comput. Human Behav.* 124:106910. doi: 10.1016/j.chb.2021.106910
- De Vicente Domínguez, A. M., Beriain Bañares, A., and Sierra Sánchez, J. (2021). Young Spanish adults and disinformation: do they identify and spread fake news and are they literate in it? *Publica* 9:2. doi: 10.3390/publications9010002
- Department for Education (2023) State of the nation 2022: Children and young people's wellbeing (Washington, DC; Hershey, Pennsylvania; London: Government Social Research)
- Diaz Ruiz, C., and Nilsson, T. (2023). Disinformation and echo chambers: how disinformation circulates on social media through identity-driven controversies. *J. Public Policy Mark.* 42, 18–35. doi: 10.1177/07439156221103852
- Diepeveen, S., and Pinet, M. (2022). User perspectives on digital literacy as a response to misinformation. *Dev. Policy Rev.* 40:e12671. doi: 10.1111/dpr.12671
- Dimock, M. (2019). Where millennials end and generation Z begins. Washington, DC: Pew Research Center.
- Du, Y. R. (2023). Personalisation, echo chambers, news literacy, and algorithmic literacy: a qualitative study of AI-powered news app users. *J. Broadcast. Electron. Media* 67, 246–273. doi: 10.1080/08838151.2023.2182787
- Duan, T. (2023). Preparing youth with media literacy for future digital civic engagement. *Proceedings of the 7th International Conference on Education and Multimedia Technology*, 315–320
- Durrant, A. C., Kirk, D. S., Moncur, W., Orzech, K. M., Taylor, R., and Trujillo Pisanty, D. (2018). Rich pictures for stakeholder dialogue: a polyphonic picture book. *Des. Stud.* 56, 122–148. doi: 10.1016/j.destud.2018.01.001
- Ecker, U. K. H., Lewandowsky, S., Cook, J., Schmid, P., Fazio, L. K., Brashier, N., et al. (2022). The psychological drivers of misinformation belief and its resistance to correction. *Nat. Rev. Psychol.* 1, 13–29. doi: 10.1038/s44159-021-00006-y
- Elgot, J., and Courea, E. (2024). Online safety act not fit for purpose after far-right riots, says Sadiq Khan. The Guardian. Available online at: <https://www.theguardian.com/media/article/2024/aug/08/online-safety-act-not-fit-for-purpose-far-right-riots-sadiq-khan> (Accessed December 22, 2024).
- Elsden, C., Chatting, D., Durrant, A. C., Garbett, A., Nissen, B., Vines, J., et al. (2017). On speculative enactments. *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 5386–5399
- Enoch, F., Johansson, P., Bright, J., and Margetts, H. Z. (2023). Tracking experiences of online harms and attitudes towards online safety interventions: findings from a

- large-scale, nationally representative survey of the British public. The Alan Turing Institute. Available online at: <https://www.ssrn.com/abstract=4416355> (Accessed December 22, 2024).
- Eysenbach, G. (2020). How to fight an Infodemic: the four pillars of Infodemic management. *J. Med. Internet Res.* 22:e21820. doi: 10.2196/21820
- Fathallah, S. (2021). The case for participatory methods in design research [medium]. Available online at: <https://medium.com/pollinator/the-case-for-participatory-methods-in-design-research-9ca23ac55cd8> (Accessed December 22, 2024).
- Foley, S., Welsh, D., Pantidi, N., Morrissey, K., Nappey, T., and McCarthy, J. (2019). Printer pals: experience-centered design to support agency for people with dementia. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–13
- Ford, C., MacKay, M., McWhirter, J., and Papadopoulos, A. (2025). Communicating about COVID-19 vaccines: a qualitative study on preferences, attitudes, and influences on Canadian post-secondary student decisions to get vaccinated. *Emerg. Adulthood* 13, 175–184. doi: 10.1177/21676968241292110
- Ford, C., MacKay, M., Thaivalappil, A., McWhirter, J., and Papadopoulos, A. (2024). COVID-19 vaccine communications on Instagram and vaccine uptake in young adults: a content assessment and public engagement analysis. *Emerg. Adulthood* 12, 224–235. doi: 10.1177/21676968231222439
- Freelon, D., and Wells, C. (2020). Disinformation as political communication. *Polit. Commun.* 37, 145–156. doi: 10.1080/10584609.2020.1723755
- Full Fact (2024) The online safety act and misinformation: what you need to know full fact. Available online at: <https://fullfact.org/policy/online-safety-act/#:~:text=The%20Online%20Safety%20Act%202023,UK%20regulator%20for%20online%20safety> (Accessed December 22, 2024).
- Gamage, D., Stomber, J., Jahanbakhsh, F., Skeet, B., and Shahi, G. K. (2022). Designing Credibility tools to combat mis/disinformation: a human-centered approach. *CHI Conference on Human Factors in Computing Systems Extended Abstracts*, 1–4
- Garg, R. (2022). Supporting the design of smart speakers to foster a sense of ownership in asian indian families. *CHI Conference on Human Factors in Computing Systems*, 1–15
- Golembewski, M., and Selby, M. (2010). Ideation decks: A card-based design ideation tool. *Proceedings of the 8th ACM Conference on Designing Interactive Systems*, 89–92.
- Gonzalez Montero, S. A., Sarria-Palacio, A. M., López Gómez, C., and Sierra Montero, J. M. (2024). Exploring collective agency: a methodological approach to becoming differently. *J. Cult. Res.* 28, 393–414. doi: 10.1080/14797585.2024.2386947
- Greenberg, S., Carpendale, S., Marquardt, N., and Buxton, B. (2012). The narrative storyboard: telling a story about use and context over time. *Interactions* 19, 64–69. doi: 10.1145/2065327.2065340
- Güldenpfennig, F., Mayer, P., Panek, P., and Fitzpatrick, G. (2019). An autonomy-perspective on the design of assistive technology experiences of people with multiple sclerosis. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–14
- Gurgun, S., Cemiloglu, D., Close, E. A., Phalp, K., Nakov, P., and Ali, R. (2024). Why do we not stand up to misinformation? Factors influencing the likelihood of challenging misinformation on social media and the role of demographics. *Technol. Soc.* 76:102444. doi: 10.1016/j.techsoc.2023.102444
- Haider, J., Söderström, K. R., Ekström, B., and Rödl, M. (2024). GPT-fabricated scientific papers on Google scholar: key features, spread, and implications for preempting evidence manipulation. *Harv. Kennedy Sch. Misinform. Rev.* doi: 10.37016/mr-2020-156
- Hainz, T., Bossert, S., and Strech, D. (2016). Collective agency and the concept of ‘public’ in public involvement: a practice-oriented analysis. *BMC Med. Ethics* 17:1. doi: 10.1186/s12910-015-0083-z
- Han, Q., and Curtis, D. R. (2021). Suspicious minds: cinematic depiction of distrust during epidemic disease outbreaks. *Med. Humanit.* 47, 248–256. doi: 10.1136/medhum-2020-011871
- Haque, M. M., Yousuf, M., Alam, A. S., Saha, P., Ahmed, S. I., and Hassan, N. (2020). Combating misinformation in Bangladesh: roles and responsibilities as perceived by journalists, fact-checkers, and users. *Proc. ACM Hum.-Comput. Interact.* 4, 1–32. doi: 10.1145/3415201
- Harrington, C., Erete, S., and Piper, A. M. (2019). Deconstructing community-based collaborative design: Towards more equitable participatory design engagements. *Proc. ACM Hum.-Comput. Interact.* 3, 1–25. doi: 10.1145/3359318
- Hassoun, A., Beacock, I., Consolvo, S., Goldberg, B., Kelley, P. G., and Russell, D. M. (2023). Practicing information sensibility: How Gen Z engages with online information. *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, 1–17
- Hope, S., and Mulhall, H. (2024). Cards on the table: critical reflections on a participatory research method. *Conjunctions* 11, 1–14. doi: 10.2478/tjcp-2023-0008
- Hopwood, K. (2023) How might vulnerable children be impacted by misinformation and disinformation online and what can parents do to limit these impacts? [internet matters]
- Humphrey, B. (2023). *Libertas Veritas: freedom and truth. The School Librarian; Wanborough* 71:26.
- Humphrecht, E., Esser, F., and Van Aelst, P. (2020). Resilience to online disinformation: a framework for cross-national comparative research. *Int. J. Press/Politics* 25, 493–516. doi: 10.1177/1940161219900126
- Hutchinson, J. W., Hutchinson, J., Vlachokyriakos, V., Kirk, D., Johnson, I. G., Evans, M., et al. (2025). Alt(ernate) CHI: Using alternate history artifacts in research. *Proceedings of the Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*, 1–10
- Insider Intelligence. (2023) Generation Z: Latest gen Z news, research, facts 2023. Available online at: <https://www.insiderintelligence.com/insights/generation-z-facts/> (Accessed July 30, 2023).
- Jabbar, M. K., Al-GBuri, M. F. A. R., and Alwan, H. S. (2024). The development of fast and slow thinking of adolescents and adults. *Social Space* 24:66–85.
- Jaffe, A. E., Graupensperger, S., Blayney, J. A., Duckworth, J. C., and Stappenbeck, C. A. (2022). The role of perceived social norms in college student vaccine hesitancy: implications for COVID-19 prevention strategies. *Vaccine* 40, 1888–1895. doi: 10.1016/j.vaccine.2022.01.038
- Jicol, C., Wan, C. H., Doling, B., Illingworth, C. H., Yoon, J., Headey, C., et al. (2021). Effects of emotion and agency on presence in virtual reality. *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 1–13
- Jimenez, N., Rodriguez-Lara, I., Tyrann, J.-R., and Wengström, E. (2018). Thinking fast, thinking badly. *Econ. Lett.* 162, 41–44. doi: 10.1016/j.econlet.2017.10.018
- Joelsson, T., and Bruno, L. (2022). Proximal or peripheral: temporality and spatiality in young people’s discourses on gender violence in Sweden. *Gender Educ.* 34, 167–182. doi: 10.1080/09540253.2020.1860199
- Joris, G., Grove, F. D., Van Damme, K., and De Marez, L. (2021). Appreciating news algorithms: examining audiences’ perceptions to different news selection mechanisms. *Digit. Journal.* 9, 589–618. doi: 10.1080/21670811.2021.1912626
- Juvalta, S., Speranza, C., Robin, D., El Maohub, Y., Krasselt, J., Dreesen, P., et al. (2023). Young people’s media use and adherence to preventive measures in the “infodemic”: is it masked by political ideology? *Soc. Sci. Med.* 317:115596. doi: 10.1016/j.socscimed.2022.115596
- Kahneman, D. (2012). *Thinking, fast and slow*. New York, NY: Penguin.
- Kalaja, P., and Rouhotie-Lyhty, M. (2021). “Autonomy and agency” in The Routledge handbook of the psychology of language learning and teaching. 1st ed, eds. González-Lloret, M., Lasagabaster, D., Verhoeven, L., and Etxeberria, A. (New York, NY: Routledge), 245–259.
- Kaluža, J. (2022). Habitual generation of filter bubbles: why is algorithmic personalisation problematic for the democratic public sphere? *Javnost Public* 29, 267–283. doi: 10.1080/13183222.2021.2003052
- Karlova, N. A. (2018) Misinformation and disinformation in online games [University of Washington]. Available online at: https://digital.lib.washington.edu/researchworks/bitstream/handle/1773/42416/Karlova_washington_0250E_18480.pdf?sequence=1&isAllowed=y (Accessed December 22, 2024).
- Kasahara, S., Nishida, J., and Lopes, P. (2019). Preemptive action: Accelerating human reaction using electrical muscle stimulation without compromising agency. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–15
- Kavada, A. (2016). Social movements and political agency in the digital age: a communication approach. *Media Commun.* 4, 8–12. doi: 10.17645/mac.v4i4.691
- Khrystenko, O. (2022). Implicit displays of emotional vulnerability: a cross-cultural analysis of “unacceptable” embarrassment-related emotions in the communication within male groups. *Open Linguist.* 8, 209–231. doi: 10.1515/opli-2022-0189
- Kim, J. W. (2018). Rumor has it: the effects of virality metrics on rumor believability and transmission on twitter. *New Media Soc.* 20, 4807–4825. doi: 10.1177/1461444818784945
- Kim, A., Moravec, P. L., and Dennis, A. R. (2019). Combating fake news on social media with source ratings: the effects of user and expert reputation ratings. *J. Manage. Inf. Syst.* 36, 931–968. doi: 10.1080/07421222.2019.1628921
- Koltay, T. (2011). The media and the literacies: media literacy, information literacy, digital literacy. *Media Cult. Soc.* 33, 211–221. doi: 10.1177/0163443710393382
- KPI (2024). *Beware of the automation paradox: The vital role of humans in automation*. Melbourne, VIC, Australia: KPI.
- Lawson, A. P., and Mayer, R. E. (2022). The power of voice to convey emotion in multimedia instructional messages. *Int. J. Artif. Intell. Educ.* 32, 971–990. doi: 10.1007/s40593-021-00282-y
- Lazar, A., Edasis, C., and Piper, A. M. (2017). Supporting people with dementia in digital social sharing. *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 2149–2162
- Lebon, G. (2008) *The crowd: The study of the popular mind*. Digireads.com
- Lee, H., and Oh, H. J. (2017). Normative mechanism of rumor dissemination on twitter. *Cyberpsychol. Behav. Soc. Netw.* 20, 164–171. doi: 10.1089/cyber.2016.0447
- Lepper, J. (2024). Digitally excluded children at greater risk of fake news, researchers warn. Children & Young People Now. Available online at: <https://www.cypnow.co.uk/content/news/digitally-excluded-children-at-greater-risk-of-fake-news-researchers-warn/#:~:text=This%20violence%20was%20sparked%20by,death%20in%20Southport%20last%20month> (Accessed December 22, 2024).
- Li, J., and Chang, X. (2023). Combating misinformation by sharing the truth: a study on the spread of fact-checks on social media. *Inf. Syst. Front.* 25, 1–15. doi: 10.1007/s10796-022-10296-z
- Li, C., Kalyanaraman, S., and Du, Y. R. (2011). Moderating effects of collectivism on customized communication: a test with tailored and targeted messages. *Asian J. Commun.* 21, 575–594. doi: 10.1080/01292986.2011.609596

- Lin, H.-C., Bruning, P. F., and Swarna, H. (2018). Using online opinion leaders to promote the hedonic and utilitarian value of products and services. *Bus. Horiz.* 61, 431–442. doi: 10.1016/j.bushor.2018.01.010
- Lin, T.-B., Li, J.-Y., Deng, F., and Lee, L. (2013). Understanding new media literacy: an explorative theoretical framework. *J. Educ. Technol. Soc.* 16, 160–170.
- Linden, C.-G. (2017). Decades of automation in the newsroom: why are there still so many jobs in journalism? *Digit. Journal.* 5, 123–140. doi: 10.1080/21670811.2016.1160791
- List, A., Grossnickle, E. M., and Alexander, P. A. (2015). Undergraduate students' justifications for source selection in a digital academic context. *J. Educ. Comput. Res.* 54, 22–61. doi: 10.1177/0735633115606659
- List, C., and Pettit, P. (2011). Group agency: The possibility, design, and status of corporate agents. Oxford UK; New York, NY, USA: Oxford University Press.
- Liu, X., Qi, L., Wang, L., and Metzger, M. J. (2023). Checking the fact-checkers: the role of source type, perceived credibility, and individual differences in fact-checking effectiveness. *Commun. Res.* 52:00936502231206419. doi: 10.1177/00936502231206419 (Accessed January 25, 2026).
- Livingstone, S. (2004). Media literacy and the challenge of new information and communication technologies. *Commun. Rev.* 7, 3–14. doi: 10.1080/10714420490280152
- Livingstone, S., Van Couvering, E., and Thumim, N. (2019). Adult media literacy. Centre for the Study of Children. Available online at: <https://www.ofcom.gov.uk/media-use-and-attitudes/media-habits-adults/adults> (accessed 25, January 2026).
- Lukoff, K., Lyngs, U., Zade, H., Liao, J. V., Choi, J., Fan, K., et al. (2021). How the design of YouTube influences user sense of agency. *Proceedings of the 2021 CHI conference on human factors in computing systems*, 1–17
- Luo, M., Hancock, J. T., and Markowitz, D. M. (2022). Credibility perceptions and detection accuracy of fake news headlines on social media: effects of truth-bias and endorsement cues. *Commun. Res.* 49, 171–195. doi: 10.1177/0093650220921321
- Lyngs, U., Lukoff, K., Csuka, L., Slovák, P., Van Kleek, M., and Shadbolt, N. (2022). The goldilocks level of support: using user reviews, ratings, and installation numbers to investigate digital self-control tools. *Int. J. Hum.-Comput. Stud.* 166:102869. doi: 10.1016/j.ijhcs.2022.102869
- MacLeod, C., Mathews, A., and Tata, P. (1986). Attentional bias in emotional disorders. *J. Abnorm. Psychol.* 95, 15–20. doi: 10.1037/0021-843X.95.1.15
- Maertens, R., Roozenbeek, J., Basol, M., and van der Linden, S. (2021). Long-term effectiveness of inoculation against misinformation: three longitudinal experiments. *J. Exp. Psychol. Appl.* 27, 1–16. doi: 10.1037/xap0000315
- Manca, S., Bocconi, S., and Gleason, B. (2021). “Think globally, act locally”: a global approach to the development of social media literacy. *Comput. Educ.* 160:104025. doi: 10.1016/j.compedu.2020.104025
- Mathiesen, K. (2006). We're all in this together: responsibility of collective agents and their members. *Midwest Stud. Philos.* 30, 240–255. doi: 10.1111/j.1475-4975.2006.00137.x
- McClure Haughey, M., Muralikumar, M. D., Wood, C. A., and Starbird, K. (2020). On the misinformation beat: Understanding the work of investigative journalists reporting on problematic information online. *Proc. ACM Hum.-Comput. Interact.*, 4, 1–22
- Mentis, H. M., Madjaroff, G., and Massey, A. K. (2019). Upside and downside risk in online security for older adults with mild cognitive impairment. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–13
- Meriläinen, M., and Ruotsalainen, M. (2024). Online disinhibition, normative hostility, and banal toxicity: Young people's negative online gaming conduct. *Soc. Media Soc.* 10:20563051241274669. doi: 10.1177/20563051241274669
- Milmo, D. (2024). Social media platforms have work to do to comply with online safety act, says Ofcom. The Guardian. Available online at: <https://www.theguardian.com/media/2024/dec/16/social-media-platforms-have-work-to-do-to-comply-with-online-safety-act-says-ofcom> (Accessed December 22, 2024).
- Mitchell, K. J., Banyard, V., and Ybarra, M. L. (2022). Are the bystanders okay? Exploring the impact of bystander behavior for self-directed violence. *J. Adolesc. Health* 70, 298–304. doi: 10.1016/j.jadohealth.2021.08.003
- Molla, R. (2021). Telegram: extremists' new favorite messaging and social media app. Vox. Available online at: <https://www.vox.com/recode/22238755/telegram-messaging-social-media-extremists> (Accessed December 22, 2024).
- Moore, R. C., and Hancock, J. T. (2022). A digital media literacy intervention for older adults improves resilience to fake news. *Sci. Rep.* 12:6008. doi: 10.1038/s41598-022-08437-0
- Moore, J. W., and Obhi, S. S. (2012). Intentional binding and the sense of agency: a review. *Conscious. Cogn.* 21, 546–561. doi: 10.1016/j.concog.2011.12.002
- Morales, E. (2024). Social media and the mediation of everyday violence: a study of Colombian young adults' experiences. *New Media Soc.* 27, 4308–4325. doi: 10.1177/14614448241236756
- Naeem, M., Ozuem, W., Howell, K., and Ranfagni, S. (2023). A step-by-step process of thematic analysis to develop a conceptual model in qualitative research. *Int J Qual Methods* 22:16094069231205789. doi: 10.1177/16094069231205789
- Nassetta, J., and Gross, K. (2020). State media warning labels can counteract the effects of foreign disinformation. *Harv. Kennedy Sch. Misinf. Rev.* doi: 10.37016/mr-2020-45
- Nazari, Z., Oruji, M., and Jamali, H. R. (2022). News consumption and behavior of young adults and the issue of fake news. *J. Inf. Sci. Theory Pract.* 10, 1–16. doi: 10.1633/JISTAP.2022.10.2.1
- Newman, N., Fletcher, R., Eddy, K., Robinson, C. T., and Nielsen, R. K. (2023). Reuters institute digital news report 2023. Oxford, UK: Reuters Institute for the Study of Journalism.
- Newman, N., Fletcher, R., Robertson, C. T., Eddy, K., and Nielsen, R. K. (2022). Reuters institute digital news report 2022. Oxford, UK: Reuters Institute for the Study of Journalism.
- Newman, N., Fletcher, R., Robertson, C. T., Ross Arguedas, A., and Nielsen, R. K. (2024). Reuters institute digital news report 2024. Oxford, UK: Reuters Institute for the Study of Journalism.
- Nurmi, J., and Jaakola, J. (2023). Losing trust: processes of vaccine hesitancy in parents' narratives. *Soc. Sci. Med.* 331:116064. doi: 10.1016/j.socscimed.2023.116064
- Ofcom (2024a) Adults' media use and attitudes report
- Ofcom. (2024b). Four in 10 UK adults encounter misinformation [Ofcom].
- Office for National Statistics (2019). Milestones: Journeying into adulthood. London, UK: ONS.
- Oh, H. J., and Lee, H. (2019). When do people Verify and Share health rumors on social media? The effects of message importance, health anxiety, and health literacy. *J. Health Commun.* 24, 837–847. doi: 10.1080/10810730.2019.1677824
- Online Safety Act 2023, UK Public General Acts (2023).
- Pal, A., Chua, A. Y. K., and Hoe-Lian Goh, D. (2020). How do users respond to online rumor rebuttals? *Comput. Hum. Behav.* 106:106243. doi: 10.1016/j.chb.2019.106243
- Paris, B., and Donovan, J. (2019). Deepfakes and cheap fakes: The manipulation of audio and visual evidence (United States of America: Data & Society). University of California. Available online at: https://datasociety.net/wp-content/uploads/2019/09/DS_Deepfakes_Cheap_FakesFinal-1-1.pdf (Accessed September 30, 2024).
- Park, P. S., Goldstein, S., O'Gara, A., Chen, M., and Hendrycks, D. (2024). AI deception: a survey of examples, risks, and potential solutions. *Patterns* 5:100988. doi: 10.1016/j.patter.2024.100988
- Parker, K., and Igielnik, R. (2020). On the cusp of adulthood and facing an uncertain future: What we know about gen Z so far. Washington, DC: Pew Research Center.
- Peng, W., Lee, H. R., and Lim, S. (2024). Leveraging Chatbots to combat health misinformation for older adults: participatory design study. *JMIR Form. Res.* 8:e60712. doi: 10.2196/60712
- Pennycook, G., and Rand, D. G. (2020). Who falls for fake news? The roles of bullshit receptivity, overclaiming, familiarity, and analytic thinking. *J. Pers.* 88, 185–200. doi: 10.1111/jopy.12476
- Peter, C., and Muth, L. (2023). Social media influencers' role in shaping political opinions and actions of young audiences. *Media Commun.* 11, 164–174. doi: 10.17645/mac.v11i3.6750
- Quintas-Froufe, N., González-Neira, A., and Fiaño-Salinas, C. (2024). Corporate policies to protect against disinformation for young audiences: the case of TikTok. *Front. Commun.* 9:1410100. doi: 10.3389/fcomm.2024.1410100
- Rezk, A. M., Simkute, A., Luger, E., Vines, J., Eldsen, C., Evans, M., et al. (2024). Agency aspirations: Understanding users' preferences and perceptions of their role in personalised news curation. *Proceedings of the CHI Conference on Human Factors in Computing Systems*, 1–16.
- Roberts, K., Dowell, A., and Nie, J.-B. (2019). Attempting rigour and replicability in the thematic analysis of qualitative research data; a case study of codebook development. *BMC Med. Res. Methodol.* 19:66. doi: 10.1186/s12874-019-0707-y
- Roozenbeek, J., and Van Der Linden, S. (2019). Fake news game confers psychological resistance against online misinformation. *Palgrave Commun.* 5:65. doi: 10.1057/s41599-019-0279-9
- Rydenfelt, H. (2022). Transforming media agency? Approaches to automation in Finnish legacy media. *New Media Soc.* 24, 2598–2613. doi: 10.1177/1461444821998705
- Santos, V. V. D. A., and Pereira, C. P. (2024). Promoting prevention to fake news through an educational software. *J. Interact. Syst.* 15, 170–193. doi: 10.5753/jis.2024.3769
- Schwartz, S. J., Hardy, S. A., Zamboanga, B. L., Meca, A., Waterman, A. S., Picariello, S., et al. (2015). Identity in young adulthood: links with mental health and risky behavior. *J. Appl. Dev. Psychol.* 36, 39–52. doi: 10.1016/j.appdev.2014.10.001
- Scott, L., Coventry, L., Cecchinato, M. E., and Warner, M. (2023). “I figured her feeling a little bit bad was worth it to not spread that kind of hate”: exploring how UK families discuss and challenge misinformation. *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, 1–15.
- Sehat, C. M., Li, R., Nie, P., Prabhakar, T., and Zhang, A. X. (2024). Misinformation as a harm: structured approaches for fact-checking prioritization. *Proc. ACM Hum.-Comput. Interact.* 8, 1–36. doi: 10.1145/3641010
- Shao, C., Ciampaglia, G. L., Varol, O., Yang, K.-C., Flammini, A., and Menczer, F. (2018). The spread of low-credibility content by social bots. *Nat. Commun.* 9:4787. doi: 10.1038/s41467-018-06930-7
- Shin, D. (2020). User perceptions of algorithmic decisions in the personalized AI system: perceptual evaluation of fairness, accountability, transparency, and explainability. *J. Broadcast. Electron. Media* 64, 541–565. doi: 10.1080/08838151.2020.1843357
- Shin, D., Zaid, B., Biocca, F., and Rasul, A. (2022). In platforms we trust? Unlocking the black-box of news algorithms through interpretable AI. *J. Broadcast. Electron. Media* 66, 235–256. doi: 10.1080/08838151.2022.2057984

- Shusas, E. (2024). Designing better credibility indicators: Understanding how emerging adults assess source credibility of misinformation identification and labeling. *Designing Interactive Systems Conference*, 41–44.
- Simpson, A. R. (2018). Changes in Young adulthood. Young adult development project. Available online at: <https://hr.mit.edu/static/worklife/youngadult/changes.html> (Accessed December 22, 2024).
- Söderberg, B. (2016) Inside the echo chamber: A qualitative study on auto-immigration internet media, political polarisation and social trust in a fragmented digital landscape [Södertörns högskola]
- Staender, A., and Humprecht, E. (2023). “Content analysis in the research field of disinformation” in *Standardisierte Inhaltsanalyse in der Kommunikationswissenschaft – Standardized content analysis in communication research*. eds. F. Oehmer-Pedrazzi, S. H. Kessler, E. Humprecht, K. Sommer and L. Castro (London, UK: Springer Fachmedien Wiesbaden), 339–348.
- Suler, J. (2004). The online disinhibition effect. *Cyberpsychol. Behav.* 7, 321–326. doi: 10.1089/1094931041291295
- Treadway, M., and Holloway, E. (2018) Young adults in transition: factors that support and hinder growth and change. Available online at: <http://www.natsap-jtsp.com/article/3225-young-adults-in-transition-factors-that-support-and-hinder-growth-and-change> (Accessed December 22, 2024).
- Ulrich, B. (1986). *Risk society: Towards a new modernity*. New Delhi: Sage.
- Valencia, S., Pavel, A., Santa Maria, J., Yu, S. (Gloria), Bigham, J. P., et al. (2020). Conversational agency in augmentative and alternative communication. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–12
- Veeriah, J. (2021). Young adults’ ability to detect fake news and their new media literacy level in the wake of the COVID-19 pandemic. *J. Content Community Commun.* 13:5168788. doi: 10.31620/JCCC.06.21/31.372
- Verwiebe, R., Buder, C., Weissmann, S., Osorio-Krauter, C., and Philipp, A. (2024). “The algorithm is like a mercurial god”: exploring content creators’ perception of algorithmic agency on YouTube. *New Media Soc.* 28:14614448241307931. doi: 10.1177/14614448241307931
- Vlachokyriakos, V., Johnson, I., Anderson, R., Claisse, C., Zhang, V., and Briggs, P. (2024). Design implications for a social and collaborative understanding of online information assessment practices, challenges and heuristics. *Proceedings of the 22nd European Conference on Computer-Supported Cooperative Work (ECSCW 2024)*. European Society for Socially Embedded Technologies, 1–19. doi: 10.48340/ecscw2024_n04
- Walther, S., and McCoy, A. (2021). US extremism on telegram: Fueling disinformation, conspiracy theories, and accelerationism. *Perspectives on Terrorism*, 15, 100–116.
- Wang, Y., McKee, M., Torbica, A., and Stuckler, D. (2019). Systematic literature review on the spread of health-related misinformation on social media. *Soc. Sci. Med.* 240:112552. doi: 10.1016/j.socscimed.2019.112552
- Wollebæk, D., Karlsen, R., Steen-Johnsen, K., and Enjolras, B. (2019). Anger, fear, and echo chambers: the emotional basis for online behavior. *Soc. Media Soc.* 5:2056305119829859. doi: 10.1177/2056305119829859
- Wszalek, J. (2017). Ethical and legal concerns associated with the comprehension of legal language and concepts. *AJOB Neurosci.* 8, 26–36. doi: 10.1080/21507740.2017.1285821
- Xiao, X., Su, Y., and Lee, D. K. L. (2021). Who consumes new media content more wisely? Examining personality factors, SNS use, and new media literacy in the era of misinformation. *Soc. Media Soc.* 7:2056305121990635. doi: 10.1177/2056305121990635
- Yamaguchi, S., Kawata, Y., Murofushi, Y., and Ota, T. (2022). The development and validation of an emotional vulnerability scale for university students. *Front. Psychol.* 13:941250. doi: 10.3389/fpsyg.2022.941250
- Yang, Y., Ta, N., Li, K., Jiao, F., Hu, B., and Li, Z. (2021). Influential factors on collective anxiety of online topic-based communities. *Front. Psychol.* 12:740065. doi: 10.3389/fpsyg.2021.740065
- Young, R., Kananovich, V., and Johnson, B. G. (2023). Young adults’ folk theories of how social media harms its users. *Mass Commun. Soc.* 26, 23–46. doi: 10.1080/15205436.2021.1970186
- Yu, W., Payton, B., Sun, M., Jia, W., and Huang, G. (2023). Toward an integrated framework for misinformation and correction sharing: a systematic review across domains. *New Media Soc.* 25, 2241–2267. doi: 10.1177/14614448221116569