frontiers | Research Topics

Technology for the greater good? The influence of (ir)responsible systems on human emotions, thinking and behavior

Edited by
Andrea Szymkowiak, Lynsay Shepherd, Martina Ziefle, Philipp Brauner and Jason Nurse

Published in
Frontiers in Computer Science
Frontiers in Psychology

ISSN 1664-8714
ISBN 978-2-8325-4819-6
DOI 10.3389/978-2-8325-4819-6

About Frontiers

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

Frontiers journal series

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

Dedication to quality

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

What are Frontiers Research Topics?

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

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

Technology for the greater good? The influence of (ir)responsible systems on human emotions, thinking and behavior

Topic editors

Andrea Szymkowiak — Abertay University, United Kingdom
Lynsay Shepherd — Abertay University, United Kingdom
Martina Ziefle — RWTH, Aachen University, Germany
Philipp Brauner — RWTH, Aachen University, Germany
Jason Nurse — University of Kent, United Kingdom

Citation

Szymkowiak, A., Shepherd, L., Ziefle, M., Brauner, P., Nurse, J., eds. (2024). *Technology for the greater good? The influence of (ir)responsible systems on human emotions, thinking and behavior*. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-8325-4819-6

Table of contents

frontiers | Frontiers in Computer Science

TYPE Editorial
PUBLISHED 13 December 2023
DOI 10.3389/fcomp.2023.1341692

OPEN ACCESS

EDITED AND REVIEWED BY
Roberto Therón,
University of Salamanca, Spain

*CORRESPONDENCE
Andrea Szymkowiak
✉ a.szymkowiak@abertay.ac.uk

RECEIVED 20 November 2023
ACCEPTED 22 November 2023
PUBLISHED 13 December 2023

CITATION
Szymkowiak A, Shepherd LA, Nurse JRC, Brauner P and Ziefle M (2023) Editorial: Technology for the greater good? The influence of (ir)responsible systems on human emotions, thinking and behavior. *Front. Comput. Sci.* 5:1341692. doi: 10.3389/fcomp.2023.1341692

Editorial: Technology for the greater good? The influence of (ir)responsible systems on human emotions, thinking and behavior

Andrea Szymkowiak[1]*, Lynsay A. Shepherd[1], Jason R. C. Nurse[2], Philipp Brauner[3] and Martina Ziefle[3]

[1]School of Design and Informatics, Abertay University, Dundee, United Kingdom, [2]School of Computing, University of Kent, Canterbury, United Kingdom, [3]Human-Computer Interaction Center, RWTH Aachen University, Aachen, Germany

KEYWORDS

human-centered technology, artificial intelligence, user experience, social media, robotics, user characteristics, security and privacy

Editorial on the Research Topic
Technology for the greater good? The influence of (ir)responsible systems on human emotions, thinking and behavior

Advanced technological systems have a tremendous impact on our lives, organizations, and societies (Stephanidis et al., 2019). We may be checking our communications on our smart phones in the morning, accessing our social media sites and operating smart systems on-the-go, while reading the latest articles based on AI-powered recommendations. Arguably every aspect of our lives is entangled with technology, be it how we communicate and interact with each other, how we entertain ourselves, how we maintain our households, safety, security, and wellbeing, how we manage our resources, or how we travel, work, or educate ourselves. Like never before, social media platforms provide a universal means of networking with others, spanning the globe, with immediate impact, and their access is only limited by cultural or political frontiers. Digital characters and robots provide us with social support and companionship, thus exceeding their traditional role of providing utilitarian value. In line with Reeves and Nass's (1996) seminal book "The Media Equation," the notion of computers as social actors continues to inspire future work, especially as digital entities appear to behave like sentient beings in increasingly sophisticated ways. Our decision-making may be influenced by recommender systems or social media, and smart systems take over tasks that we ourselves performed in the past. The ubiquitous presence of technology systems affects societies in many ways, raising interesting philosophical perspectives (Van de Poel, 2020).

Against this background, research on human-technology interactions has barely kept pace, resulting in numerous articles published concurrently with new technological inventions. Technology use is inextricably linked with some form of user experience. However, the consequences of human-technology interactions are not always clearcut. To the extent that we shape technology, it shapes us, and often in unanticipated ways. AI, due to its malleable nature, can be compromised, e.g., as in the case of Microsoft's chat bot "Tay," which produced inflammatory posts within 1 day of usage (Neff and Nagy, 2016). There is well-documented evidence regarding social media use and its (adverse) effects on psychological wellbeing (Twenge, 2019). It will be of interest to observe whether/how

law and policy makers regulate the design of choice architectures and monitor technology-induced incidents in the future to prevent harm. With the proliferation of such technologies, it can be argued that new challenges, but also opportunities, will arise for human (and non-human) users of such systems.

This Research Topic on "*Technology for the greater good?*" comprises a collection of papers exploring novel work in human-technology interaction, with the aim of identifying new challenge spaces and topics. The scope of the Research Topic, given the volume of innovations, is necessarily non-exhaustive, but aims to provide an overview of pertinent issues that affect current society and individuals.

The first group of studies explores perceived privacy and security aspects of technology. Belen-Saglam et al. studied disclosure of sensitive data and findings suggest that the use of conversational agents detrimentally affected disclosure in the health domain, but less so in the financial domain. Hildebrandt et al. explored users' privacy concerns in mobile forensics, with users showing a preference for the release of less personal data, such as geo-spatial data over more personal data such as photographs and favoring automated rather than human evaluations. Finally, Brauner et al. examined public perceptions on the use of AI, reporting that cybersecurity threats were perceived to be highly likely and least liked. People scoring higher on dispositional trust had more favorable views of AI compared to people with lower trust. The findings highlight the intricacies of user decision-making and user acceptance in relation to handling sensitive material, which should be considered by creators of technology.

The second group of contributions investigates user experiences or behavior when interacting with robots or robot process automation. Employing a lab-based study, Maalouly et al. reported that users were more altruistic toward a humanoid robot after sustained conversation, which suggests that anthropomorphic technology can elicit pro-social behavior. Filgueiras et al. showed that, while multi-faceted, the user experience after prolonged use with robot process automation reflected user acceptance and adoption, especially where automation provided utilitarian value to the user.

The final set of articles highlights technology interactions with people's dispositions, states, and experiences. In relation to procrastination, Sümer and Büttner reported that higher boredom proneness, lower self-control and lower perseverance were predictive of different types of online procrastination. With a focus on social media and mental health, Scarpulla et al. showed that more active social media use was associated with increased anxiety and stress as well as poorer emotion recognition skill, while passive social media use was not associated with these variables. Protzko and Schooler demonstrated that people were more inclined to view technological-societal shifts as corruptive of today's youth if they had not encountered this technology themselves during their formative years, pointing to the role that personal experience plays for beliefs about society and the technology it is exposed to. Xie and Liu demonstrated that trust in social media platforms relates to perceived information quality, perceived privacy, a sense of social belonging and sense of self-esteem, and positive emotion. The work by Kaminger et al. revealed that dispositional gratitude can act as a protective factor when using social media by moderating the relationship of social comparison, and malicious and general envy. The experiences of and interactions with human-centered technology are multi-faceted and give rise to equally diverse research findings.

In conclusion, we thank all involved in the preparation of this Research Topic, contributing from various disciplines, countries, and contexts. The contributions underscore the importance of considering user perceptions and experiences as pivotal factors in steering future human-technology innovations. We hope that these developments ultimately contribute to the creation of systems that assist and benefit their users and society, removing the question mark in the title of this Research Topic—in other words, in technology for the greater good.

Author contributions

AS: Conceptualization, Writing – original draft. LS: Writing – review & editing. JN: Writing – review & editing. PB: Writing – review & editing. MZ: Writing – review & editing.

Funding

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

Conflict of interest

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

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Publisher's note

References

Neff, G., and Nagy, P. (2016). Talking to bots: symbiotic agency and the case of Tay. *Int. J. Commun.* 10, 4915–4931.

Reeves, B., and Nass, C. (1996). *The Media Equation: How People Treat Computers, Television, and New Media like Real People and Places*. Cambridge: Cambridge University Press.

Stephanidis, C., Salvendy, G., Antona, M., Chen, J. Y., Dong, J., Duffy, V. G., et al. (2019). Seven HCI grand challenges. *Int. J. Human–Comput. Interact.* 35, 1229–1269. doi: 10.1080/10447318.2019.1619259

Twenge, J. M. (2019). More time on technology, less happiness? Associations between digital-media use and psychological well-being. *Curr. Dir. Psychol. Sci.* 28, 372–379. doi: 10.1177/0963721419838244

Van de Poel, I. (2020). Three philosophical perspectives on the relation between technology and society, and how they affect the current debate about artificial intelligence. *Human Affairs* 30, 499–511. doi: 10.1515/humaff-2020-0042

frontiers | Frontiers in Computer Science

ORIGINAL RESEARCH
published: 30 June 2022
doi: 10.3389/fcomp.2022.908245

An Investigation Into the Sensitivity of Personal Information and Implications for Disclosure: A UK Perspective

Rahime Belen-Saglam[1], *Jason R. C. Nurse*[1]* *and Duncan Hodges*[2]

[1] School of Computing, University of Kent, Canterbury, United Kingdom, [2] Centre for Electronic Warfare, Information and Cyber, Cranfield University, Defence Academy of the United Kingdom, Shrivenham, United Kingdom

The perceived sensitivity of information is a crucial factor in both security and privacy concerns and the behaviors of individuals. Furthermore, such perceptions motivate how people disclose and share information with others. We study this topic by using an online questionnaire where a representative sample of 491 British citizens rated the sensitivity of different data items in a variety of scenarios. The sensitivity evaluations revealed in this study are compared to prior results from the US, Brazil and Germany, allowing us to examine the impact of culture. In addition to discovering similarities across cultures, we also identify new factors overlooked in the current research, including concerns about reactions from others, personal safety or mental health and finally, consequences of disclosure on others. We also highlight a difference between the regulatory perspective and the citizen perspective on information sensitivity. We then operationalized this understanding within several example use-cases exploring disclosures in the healthcare and finance industry, two areas where security is paramount. We explored the disclosures being made through two different interaction means: directly to a human or chatbot mediated (given that an increasing amount of personal data is shared with these agents in industry). We also explored the effect of anonymity in these contexts. Participants showed a significant reluctance to disclose information they considered "irrelevant" or "out of context" information disregarding other factors such as interaction means or anonymity. We also observed that chatbots proved detrimental to eliciting sensitive disclosures in the healthcare domain; however, within the finance domain, there was less effect. This article's findings provide new insights for those developing online systems intended to elicit sensitive personal information from users.

Keywords: personal information disclosure, information sensitivity, privacy, chatbots, conversational agents, artificial intelligence, personal information

OPEN ACCESS

Edited by:
Gualtiero Volpe,
University of Genoa, Italy

Reviewed by:
Thanh Van Do,
Telenor, Norway
Dan-Cristian Dabija,
Babeş-Bolyai University, Romania

***Correspondence:**
Jason R. C. Nurse
j.r.c.nurse@kent.ac.uk

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

Received: *30 March 2022*
Accepted: *08 June 2022*
Published: *30 June 2022*

Citation:
Belen-Saglam R, Nurse JRC and Hodges D (2022) An Investigation Into the Sensitivity of Personal Information and Implications for Disclosure: A UK Perspective. Front. Comput. Sci. 4:908245. doi: 10.3389/fcomp.2022.908245

1. INTRODUCTION

The internet has enabled people throughout the world to connect with each other in ways that previously would have been considered unimaginable. To enable such interactions, individuals are often required to share various types of information and this can in turn lead to privacy concerns about how their personal information is stored, processed and disclosed to others.

From research, we know that a user's privacy concerns and their willingness to disclose information are affected by the perceived sensitivity of that information (Markos et al., 2018).

However, it is vague and open to debate as to how "sensitive" information may be categorized. A risk-oriented definition is adopted by some studies in the literature as seen in the EU's General Data Protection Regulation (GDPR) (European Parliament, 2016) which defines sensitive information as follows:

> *Personal data which are, by their nature, particularly sensitive in relation to fundamental rights and freedoms merit specific protection as the context of their processing could create significant risks to the fundamental rights and freedoms.*

However, several other dimensions are also introduced to explain how users perceive sensitivity including: perceived risk, possibility of harm or public availability of data can lead information to be perceived as sensitive (Ohm, 2014; Rumbold and Pierscionek, 2018). In addition to studies which explore the factors leading to a high perceived sensitivity, it is possible to report two other research themes in this area. Firstly, studies that report the perceived sensitivity of different data items at granular levels or in different usage contexts (Markos et al., 2017; Milne et al., 2017; Schomakers et al., 2019; Belen Sağlam et al., 2022). Secondly, studies which investigate the relationship between information sensitivity and disclosure (Treiblmaier and Chong, 2013; Bansal et al., 2016; Wadle et al., 2019; Aiello et al., 2020; Belen Sağlam and Nurse, 2020).

This research aims to provide a UK perspective on the research areas identified above, a problem that is missing in existing literature. To the best of our knowledge, there is also no study that synthesizes findings associated with the factors that lead certain information to be considered sensitive, sensitivity ratings of different personal data items and the comfort felt while disclosing them under different conditions. Therefore, we formulated our research question as follows: "What are the perspectives of British citizens regarding the sensitivity of the information and the impact of different factors on the disclosure of personal information?." To answer this research question and provide key related insights into this issue, the following research objectives (RO) are defined:

- RO1: Identify the main factors that lead British citizens to regard certain information as sensitive.
- RO2: Explore the levels of sensitivity associated with the different personal data items
- RO3: Explore the impact of user factors on levels of sensitivity of the different personal data items.
- RO4: Explore if there is an international consensus on the level of sensitivity of the personal data items (comparing Germany, the US, Brazil and the UK).
- RO5: Determine the impact of context/situation (specifically finance or health domains) on an individual's level of comfort in disclosinginformation.
- RO6: Determine the impact of interaction means (human or chatbot) while sharing personal information on individual's level of comfort in disclosing information.
- RO7: Determine the impact of anonymity (identified or anonymous) on individual's level of comfort in disclosing information.

Through this research, we contribute to the literature on information sensitivity and disclosure in three novel ways:

1. We provide insights into the factors that lead to certain information being considered sensitive and provide a UK perspective on these debates.
2. We provide sensitivity ratings of different data items for UK citizens and explore the international consensus on data sensitivity. Those findings can further help to inform discussions on the process of cross-national data flows.
3. We empirically investigate the impact of demographic characteristics, anonymity, context (health and finance), and interaction means (human or chatbot) on information sensitivity and comfort to provide information.

Our findings, therefore, can also contribute to an understanding of how to design inclusive information systems when sensitive disclosures are required. The assumption we make in this study is that comfort is inversely related to sensitivity; i.e., the more comfortable an individual is in sharing some personal information, the less sensitive that information is perceived to be, this is consistent with prior work (e.g., Ackerman et al., 1999).

The remainder of this paper is structured as follows. The Literature Review section summarizes the literature relevant to our research question. We present our methodology in the Research Methodology section and following this, we present our descriptive results in Results section. We critically reflect on and consider our findings in the Discussions section, as well as highlighting the implications for research and practice. The paper closes with a discussion of the limitations of the research and future plans.

2. LITERATURE REVIEW

This section summarizes the relevant literature underpinning this research in following four sub-categories.

2.1. What Makes Information Sensitive?

A fundamental challenge for protecting personal information is first defining how it can be conceptualized and categorized. While there are several different opinions in the literature about how sensitive personal information may be defined, regulatory frameworks can provide a robust foundation. The European General Data Protection Regulation (GDPR) considers personal data sensitive if it reveals a racial or ethnic origin, political opinions, religious or philosophical beliefs, trade union membership, data concerning health, sex life and sexual orientation. In addition to these data types, genetic data and biometric data also fall into this category. The GDPR covers those data items in a special category defined as *"data that requires specific protection as the context of their processing could create significant risks to an individual's fundamental rights and freedoms"* (European Parliament, 2016).

One notable study on sensitive information, Ohm (2014) aimed to understand what makes information sensitive and focused on a list of categories of information that have been legally treated as sensitive, primarily from the United States. This list of sensitive categories was then employed to

infer the characteristics of information types that result in it being considered sensitive. In brief, four factors were reported when assessing whether a given piece of information seems sensitive: the possibility of harm, probability of harm, presence of a confidential relationship, and whether the risk reflects majoritarian concerns.

A schema has been proposed for assessing data categories to guide the relative sensitivities of different types of personal information (Rumbold and Pierscionek, 2018). The paper explores several factors that influence the perception of personal data as sensitive, including the public availability of data, the context of the data use and its potential to identify individuals. Contrary to popular belief, researchers stated that data publicly observable is not necessarily non-sensitive data (Rumbold and Pierscionek, 2018). The potential of certain information being used to infer new information when aggregated with others is another factor leading to a perception of sensitivity. Several other issues, such as the risk of re-identification, automated profiling, behavioral tracking and trustworthiness of the person/system with whom the data is shared, are also given as potential problems to affect sensitivity evaluation of particular information types. The massive increase in sensors associated the internet-of-things (IoT) devices (e.g., sensor data, or heart-rate data from wearable devices) within the medical domain has increased the amount of health data collected from citizens. This has raised the risk of third party data access such as health professionals or even insurance companies (Levallois-Barth and Zylberberg, 2017). Sharing data with third parties may increase the risk of discrimination and also make it possible to infer the prevalence of certain pathologies. Therefore, Levallois-Barth and Zylberberg (2017) claim that even though those data items may not be potentially sensitive when considered in isolation, sensitivity evaluations may change in the future. However, surprisingly, Kim et al. (2019) revealed that within healthcare, sensitivity has no statistically significant impact on the willingness to provide privacy information even though it significantly influences the perceived privacy risk. Those conflicting findings highlight some of the challenges in sensitivity evaluations and disclosure which will be explained further in Section 2.3.

Finally, the nature of the technology also has an impact on the sensitivity evaluations and data storage decisions accordingly. For instance, due to it's immutable nature which prevents data being changed, Kolan et al. (2020) argued that personal medical data should not be stored directly on public blockchain systems. This was confirmed by Zheng et al. (2018) who also preferred not to store health information in blockchain in their proposed solution. Based on that, it can be argued that the concerns regarding the use of data in the future shapes the sensitivity evaluations of personal data.

2.2. What Types of Information Are Perceived as Sensitive?

In addition to the studies that explore the factors leading individuals to perceive certain information as sensitive, studies have also categorized data types according to the perceived sensitivity.

In one of those studies researchers identified two clusters of information that were considered more sensitive: secure identifiers (e.g., social security number) and financial information (e.g., financial accounts and credit card numbers). It is noted that basic demographics (e.g., gender, birth date) and personal "preferences" (e.g., religion, political affiliation) were seen as less sensitive by the survey respondents (Milne et al., 2017).

Another study by Markos et al. (2017), used a cross-national survey between consumers in the United States and Brazil to explore the cultural differences in the perception of sensitivity. The authors examined 42 information items concluding that US consumers generally rated information as more sensitive and were less willing to provide information to others than their Brazilian counterparts. Financial information and identifiers were observed to have the highest perceived sensitivity with security codes and passwords, financial account numbers, credit card numbers, or formal identifiers such as social security number and driving license number appeared in a cluster of highly sensitive data.

A similar study has been conducted that provided a German citizen perspective on information sensitivity (Schomakers et al., 2019). Researchers compared their results with the results from the US and Brazil (Markos et al., 2017; Milne et al., 2017) and noted that, on average, the perceptions of information sensitivity of German citizens lies between that of US and Brazilian citizens. Cluster analysis revealed that similar data items were considered highly sensitive by the three countries except that German citizens considered the credit score to appear in a medium-sensitive cluster whilst US and Brazilian citizens considered this to be in a higher-sensitivity cluster. However, in general, German citizens were reported to perceive passwords as most sensitive, followed by identifiers such as financial account numbers, passport numbers or fingerprints.

In addition to those studies that focus on general items of information, some researchers focused on specific information domains. For example, Bansal et al. (2010) focused on health information and the role of individual differences on perceived information sensitivity and disclosure in this domain. Meanwhile, Ioannou et al. (2020) focused on travel providers and their customers' privacy concerns when sharing biometric and behavioral data and the impact of these concerns on the willingness to share this data. This study highlighted the context-dependence of privacy preferences. It is reported that although travelers worry about the privacy of their data, they are still willing to share their data, and the disclosure decision is dependent upon expected benefits rather than privacy concerns. Confirming the "privacy paradox" (Norberg et al., 2007), it was found that there was no link between privacy concerns and willingness to share biometric information and that expected benefits outweigh privacy concerns in the privacy decisions made by travelers.

Research has also examined attitudes toward sharing PII and non-PII (anonymous) data (Markos et al., 2018); they differentiated the information that was already public, hypothesizing that items associated with the "private-self" are perceived as more sensitive than public-self items. Their

results demonstrated that some anonymous information like diary/journal entries, hygiene habits, home information, and GPS location are considered sensitive and even more sensitive than PII, conflicting slightly with the general societal interpretation and legislative focus. More expectedly, they identified that private-self information items were perceived as more sensitive than public-self items.

2.3. When Do We Disclose More?

There are multiple debates regarding personal information disclosure in the literature, some of which consider data sensitivity and other factors such as the perception of benefit. For instance, research has found that people are more willing to disclose when their human needs such as health or security are fulfilled (Wadle et al., 2019); thus, explaining the impact of expected benefits on information disclosure.

Conversely other research proposed that the perceived privacy risks play a more significant role than the expected benefits (Keith et al., 2013). The difference in their results was explained by the high degree of realism they provided in their experiments, where participants were given a real app that dynamically showed actual data.

In another recent study, perceived privacy risks were argued to significantly reduce the intention to disclose information and the disclosure behavior, whilst privacy concerns were reported to affect disclosure intention but not the actual information disclosure behavior (Yu et al., 2020).

The impact of personal differences has also been studied; for example, less healthy individuals were more concerned about disclosing their health information arguably due to the risk of their status on employment opportunities or social standing (Bansal et al., 2016). This finding confirms previous studies by Treiblmaier and Chong (2013) who demonstrated that a higher level of perceived risk leads to a lower level of willingness to disclose personal information. The same research examined the role of trust in information disclosure and reported that the direct influence of trust in the Internet (as a communication media) is statistically insignificant. However, the trust of an online vendor (the ultimate receiver of the information) impacts the willingness to disclose.

It has also been shown that the perceived fairness of a data request also impacts personal information disclosure (Malheiros et al., 2013). The "fairness" of a data request describes the individual's belief that data being collected will be used for the purpose communicated by the data receiver and in an ethical manner. The study revealed that when participants saw a disconnect between the disclosures they were asked to make and the specified purpose of the disclosure, they consider it unfair and opted not to disclose.

The impact of anonymity has also been studied in a recent study (Schomakers et al., 2020) that reported that the critical element of online privacy and privacy in data sharing is the protection of the identity, and thus, anonymity. The most substantial effect associated with data sharing was the anonymisation level, followed by the type of data (how sensitive it is) and how much the person with whom the information is shared is trusted. It was reported that when the participants can understand why the data is useful to the receiver, they are more willing to provide data. Benefits for the self or the society are also reported as important aspects while deciding to share data. It is clear that when it comes to PII, sensitivity plays a greater role in willingness to disclose than it has for anonymous information, i.e., information that is not personally identifiable (Markos et al., 2018).

2.4. How May Non-human Agents Impact Disclosure?

A chatbot is an application created to automate tasks and imitates a real conversation with a human in their natural language (whether spoken or through a textual interface). Today, conversational agents are used in various industries, including finance and health care. In these applications, the collection of personal information is essential to provide an effective service. Consequently, research has focused on disclosing information to chatbots and the modulating factors that enable or degrade disclosure. In one of those studies, it was concluded that users disclose as much to chatbots as they would to humans (Ho et al., 2018), resulting in similar disclosure processes and outcomes. The researchers added that relatively neutral questions might not make a difference between chatbots and humans, and when asked a question that may be embarrassing and might result in negative evaluation, users were also found to respond with more disclosure intimacy to a chatbot than a human.

Another study highlighted a similar issue and noted that individuals tended to talk more freely with a chatbot, without perceiving they were being judged or making the chatbot bored of listening to them (Bjaaland and Brandtzaeg, 2018). Accessibility and anonymity are given as other characteristics of chatbots that encourage self-disclosure. "Icebreaker questions" (e.g., "how are you doing?," "how is the weather?") or human-like fillers (e.g., "um," "ahh") are also reported to lead to more effective communication and a sense of a shared experience (Bhakta et al., 2014; Bell et al., 2019).

Other research has considered the importance of context and investigated the effects of socio-emotional features on the intention to use chatbots (Ng et al., 2020). While a preference for a technical and mechanical chatbot for financially sensitive information was identified, no significant differences were observed in the disclosure of socially attributed items (such as name, date-of-birth and address) between the chatbots with and without socio-emotional traits.

The lack of coherence in the scope of the studies that investigate the impact of employing chatbots on information disclosure has encouraged us to design this study. We systematically investigate the comfort in disclosing sensitive information to a chatbot, varying the context of the domain and the sensitivity levels of data items. We aim to present a rigorous and systematic understanding of the impact on information disclosures from conversational agents.

3. RESEARCH METHODOLOGY

In order to answer our research question and achieve the individual research objectives, a rigorous methodology was defined, this was oriented around an online questionnaire

and robust qualitative and quantitative data analysis. The questionnaire engaged a sample of 500 British participants and critically explored the topic of information sensitivity. We opted for a questionnaire (e.g., instead of interviews or focus groups) to reach a census representative sample of UK citizens. The questionnaire design (i.e., questions asked, sequence of questions) and subsequent data analysis techniques were composed specifically to allow us to address each research objective, and address the research question. In what follows, we explain the questionnaire design, present the participant recruitment strategy, and detail the techniques used to analyse the data gathered.

3.1. Questionnaire Design

The questionnaire was implemented on the Survey Monkey platform, and participants were asked to respond to questions posed across five sections. First, we posed questions to collect informed consent from participants. In the second section, demographic characteristics of the participants (age group, gender, and educational level) were gathered. Having gathered this biographic information, the next sections were closely associated with the research objectives. The third section targeted RO1 specifically and therefore asked participants for the reasons or factors that might lead them to consider certain personal information more sensitive than other personal information. This was presented as an open-ended question to allow participants to present any factors they viewed appropriate.

The fourth section asked participants questions about the sensitivity of a range of personal data items. These questions provide the basis for achieving RO2 (i.e., exploring the levels of sensitivity of the different personal data items), RO3 (i.e., exploring the impact of user factors on sensitivity of the different personal data items) and RO4 [i.e., enabling a comparison of British citizens sensitivity perceptions with perceptions from citizens from the US, Brazil and Germany (Markos et al., 2017; Schomakers et al., 2019)].

To determine the data items for our study, we decided to use data items covered in existing studies as a basis and enrich those lists in accordance with our research objectives. Some of the original data items by Markos et al. (2017) and Schomakers et al. (2019) were not appropriate for our scenarios and therefore were eliminated, for example: DNA profile, fingerprint, digital signature or browsing history are not easily shared with chatbots due to their nature. We paid particular attention to the differences in the sensitivity classification of Schomakers et al. (2019) to that of Markos et al. (2017). We included the data items that were assigned different sensitivity levels between those two studies. We also expanded our list with data items considered sensitive by the GDPR or any data protection acts of EU countries, the US, China and the UK. These regulations were reviewed, and any data items that were identified as requiring extra controls or given as "special categories" were added to our list.

The complete list of data items is in **Table 1**. In order to better understand these data items within the context of the domains we considered (health and finance), these data items were manually categorized as either General data items, Health-related information, or Financial information.

TABLE 1 | The full list of data items used in the study.

Category	Data item
General data items	Passwords, Passport Number, Formal Identification Number, IP Address, Private Phone Number, Current Location, Home Address, Criminal Records, Face Picture, Online Dating Activities, Sex Life, Sexual Orientation, Email Address, Social Network Profile, License Plate Number, Shopping habits, Political Affiliation, Weight, Mother's Maiden Name, Post Code, Place Of Birth, Number Of Children, Religion, Height, Hair Color, Name Of Pet, Trade Union Membership, Social Welfare Needs, Racial or Ethnic Origin, Full Name, Education Records, Date of Birth, Citizenship, Marital Status, Gender
Health Information	Alcohol Consumption, Smoking Habits, Substance Abuse Conditions, Mental Health, HIV and/or other sexually transmitted diseases, Medical Diagnoses, Chronic Diseases
Financial Information	Credit Card Number, Credit Score, Income Level, Occupation, Bank Account Credentials

TABLE 2 | Scenarios used in the study.

ID	Interaction means	Context	Anonymity
S1	Person	Health	Anonymous
S2	Person	Finance	Anonymous
S3	Person	Health	Identified
S4	Person	Finance	Identified
S5	Chatbot	Health	Anonymous
S6	Chatbot	Finance	Anonymous

To examine participants' opinions on the sensitivity of these 40 data items, participants were asked to rank each data item on a 6-point symmetric Likert scale which ranged from "not sensitive at all" (1) to "very sensitive" (6). Throughout the study, we used a 6-point scale as done by Schomakers et al. (2019) to enable a direct comparison between nationalities. A 6-point scale has also been shown to avoid overloading the participants' discrimination abilities (Lozano et al., 2008). For the fifth and final section of the questionnaire, a set of questions was posed to assess the effects of three variables, i.e., identification (anonymous or identified), context (finance or health) and interaction means (a human or chatbot), on the comfort in disclosing personal information (RO5-7); thus, was a 2 x 2 x 2 factorial design. Participants were asked to rate their comfort level while disclosing particular data items in each of the scenarios summarized below in **Table 2**. For example, in scenario 1 (S1) the question was given as follows: "Assume that you are speaking to a person on an online health service website where you do not need to identify yourself (i.e., you can be anonymous). How comfortable would you feel disclosing (i.e., sharing) the personal information listed below?." Comfort levels were assessed again on a 6-point Likert scale ranging from 1 "Not comfortable at all" to 6 "Very comfortable."

TABLE 3 | Reduced set of 20 data items used in the final stage of the study.

Category	Data item
General data information	GPS Location, Criminal Records, Sex Life, Social Network Profile, License Plate Number, Political Affiliation, Mother's Maiden Name, Religion, Trade Union Membership, Racial or Ethnic Origin
Health information	Alcohol Consumption, Mental Health, HIV and/or other sexually transmitted diseases, Medical Diagnosis, Chronic Diseases
Finance information	Credit Card Number, Credit Score, Income Level, Occupation, Bank Account Credentials

In order to reduce the possible overload of participants, two scenarios have been eliminated from the study. These would be S7 and S8 to complete the 2 x 2 x 2 design where participants would be asked to disclose personal information to a chatbot where they needed to identify themselves. When piloting the study, it became apparent that the quality of the responses was significantly reduced beyond six scenarios. This pragmatic decision allowed us to focus on the six scenarios which would supply the most value to practitioners.

To determine the data items to use for this final part of the questionnaire, we abridged the original list of data items and selected 20 items; ten were general data items, five were health related, and five were finance related. This abridging was another pragmatic choice to reduce the load on our participants whilst still delivering a solid evidence base for practitioners. While shortening the list, we retained data items that are frequently subject to debates in the literature. Personal identifiers, data items in the special category of the GDPR or personal information related to health and finance were maintained in this list for this reason (see **Table 3**).

We included six attention checking questions to ensure the quality of our data. The scenarios in the second step were randomized in the questionnaire software to avoid any sequence bias. The data items (i.e., the lists of 40 and 20 items) in the questions were also randomized for the same purposes. The study has been reviewed and ethically approved by the Research Ethics and Governance Department of University of Kent and Cranfield University Research Ethics Committee.

3.2. Participants

Participants were recruited using Prolific in order to reach a census representative sample of UK citizens. Since this study's ultimate goal is to understand UK citizens' perspective, it was essential to gather responses from a representative set of the public. This platform was also selected since it has good quality and reproducibility compared to other crowdsourcing platforms (Peer et al., 2017).

Before running our questionnaire, we conducted a pilot study with 50 participants to ensure that the questionnaire design and time limits were appropriate and usable for the intended/target audience. We then released the complete questionnaire on a sample of 500 participants (i.e., representative of the UK population based on age, sex and ethnicity), paying £8.72 per hour, which is at least the UK minimum wage. In total, the questionnaire took 15 min to complete.

TABLE 4 | Demographic profile of participants.

Age	18-24	10.4%
	25-34	19.2%
	35-44	15.9%
	45-54	18.9%
	55-Over	35.6%
Gender	Female	50.3%
	Male	49.7%
Education	GCSE	15.5%
	A-level or equivalent	28.1%
	Undergraduate degree	34.4%
	Postgraduate degree	18.7%
	Doctorate	3.3%

GCSEs are the qualifications taken in years 10 and 11 of secondary school in the UK. A-levels are a subject-based qualification offered by the educational bodies in the UK to students completing secondary or pre-university education.

From the 500 responses gathered, nine participants failed more than one attention question and thus were excluded from the data analysis. We present the demographics of the final 491 participants in **Table 4**.

3.3. Data Analysis

To analyse the data gathered, we used techniques most appropriate for the respective question set (see **Figure 1**). After collecting consent and demographic characteristics of the participants at the beginning of the questionnaire, in the first step, to achieve RO1 we asked reasons or factors that lead participants to consider certain personal information as more sensitive than other personal information. We used thematic analysis to analyse this qualitative data (Braun and Clarke, 2006). Firstly, brief labels (codes) were produced for each response, and when all data had been initially coded, themes were identified, grouping responses with similar codes into the same category. Finally, the themes were reviewed to check whether the candidate themes appeared to form a coherent pattern.

The analysis conducted to achieve RO2 was descriptive and we ordered the data items by computing their average sensitivity ratings. For RO3, we built proportional-odds logistic regression models for each data type to model the effects of age, gender and education. This modeling approach allows us to build a model that predicts a particular participant's probability of giving a data item a particular sensitivity rating based on their age, gender, and education level. By exploring these model coefficients, we can gain insight into the effects of these variables on how comfortable people are disclosing sensitive information.

To achieve RO4, we used hierarchical cluster analysis (Bridges, 1966) to group data types based on their perceived sensitivity. Initially, each data item is assigned to an individual cluster before iterating through the data items and at each stage merging the two most similar clusters, continuing until there is one remaining cluster. At each iteration, the distance between clusters is recalculated using the Lance-Williams dissimilarity

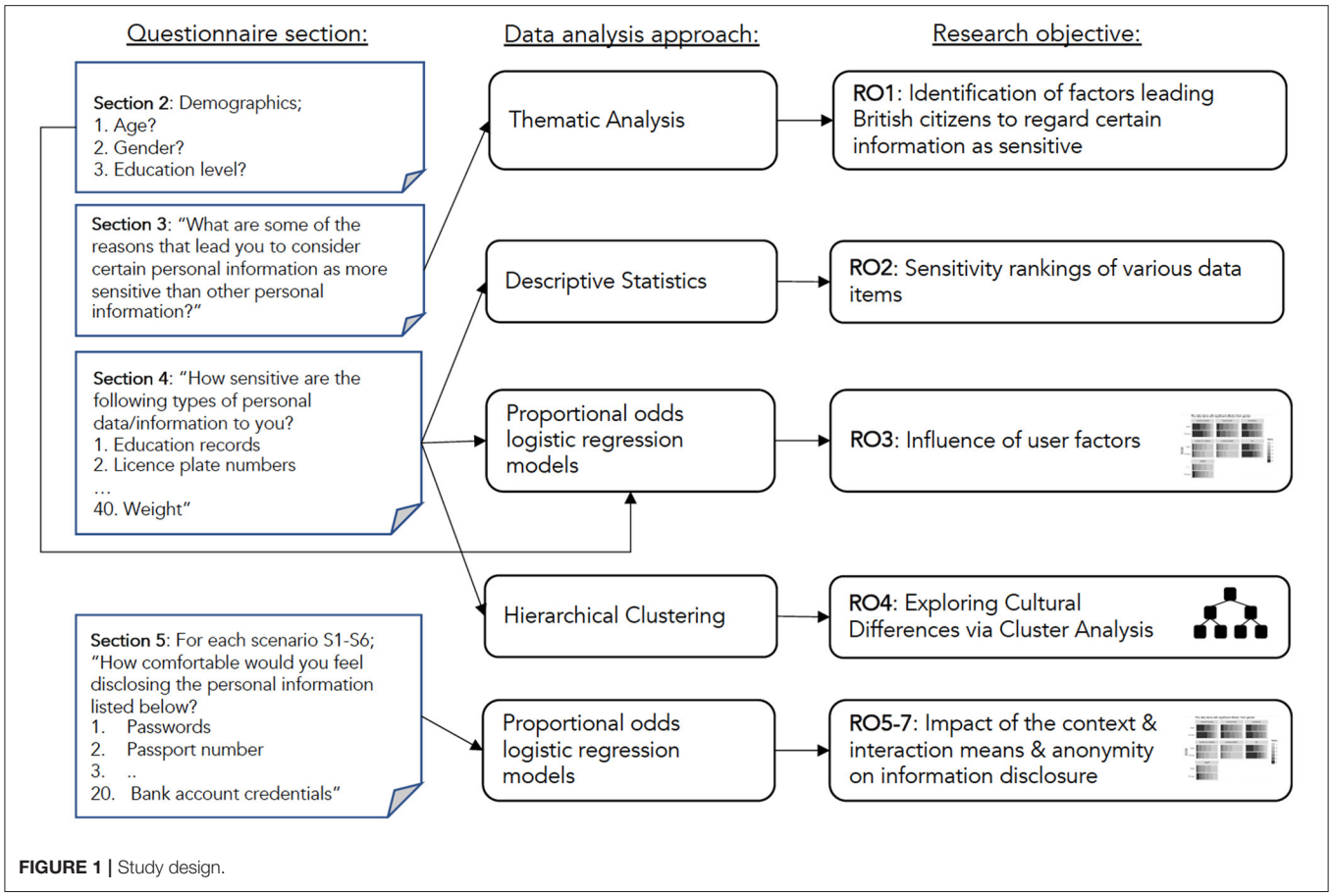


FIGURE 1 | Study design.

(Murtagh and Contreras, 2012). This clustering allowed us to build a tree diagram where the data items viewed as being of similar sensitivity are placed on close together branches.

Finally, for Research Objectives 5 to 7 we used proportional-odds logistic regression modeling to analyse the effects of anonymity, context and interaction means, using these three variables to predict the comfort level while disclosing personal information.

4. RESULTS

This section describes the results from both the open-ended qualitative question and the quantitative results from the Likert scale questions. Further discussion of the results is explored in Section 5.

4.1. RO1: Identification of Factors Leading British Citizens to Regard Certain Information Sensitive

As mentioned previously, we asked our participants an open-ended question regarding the factors that lead them to consider a data item to be sensitive. A thematic analysis of the responses led to several factors being identified. These included some of the factors reported in the literature, such as the risk of harm, trust of interaction means, public availability of data, context, and identification. However, we identified several other areas that have been overlooked or not dealt with comprehensively. These new themes included concerns regarding the reactions from the listener, concerns regarding personal safety or mental health, consequences of disclosure on beloved ones or careers, or concerns regarding sharing information about others such as family members or friends.

The complete set of themes and codes are presented in **Table 5** with the number of responses related to each theme and code. These summaries provide a useful indicator of the themes emerging from the study and the popularity of each theme.

In the remainder of this section, we provide details of the most pertinent themes emerging from our study. The names of the themes and the codes under themes are written in italics.

4.1.1. Privacy Concerns

Privacy concerns expressed by the participants while evaluating the sensitivity level of information often focused on *identity theft*. In our study, 35 participants expressed their concerns in a finance context where credentials or some other identifiers were given as examples to sensitive personal information due to their potential exploitation for identity fraud. Identifiers or other information used to identify individuals when used together were also considered sensitive by several participants even if identity theft was not explicitly mentioned. For some participants, it was

TABLE 5 | Thematic analysis of what makes data sensitive.

Themes	Codes
Privacy (181)	Identity (64), Private information (45), Identity theft (35), Access to more (18), Third party sharing (9), Personal life (5), Tracing (5)
Context (135)	Finance (80), Health (55)
Financial Problems (100)	Risk of fraud (69), Financial loss (18), Impact on career (12), Financial exposure (1)
Reactions (84)	Embarrassment (31), Discrimination (17), Judgement (15), Reputational harm (12), Cultural conditioning (5), Reactions in general (4)
Consequence of disclosure on me (84)	Personal security (18), Misuse (18), Harm (18), Personal safety (8), Risk of crime (7), Mental Health (6), Legal issues (3), Harassment (2), Cost & Benefit (1)
Nature of information (43)	Relevance (17), Public Availability (10), Information of others (7), Value (5), Group (2), Stability (1), Delicacy (1)
Interaction means (26)	Concerns regarding the recipient (20), Trust (6)
Consequence of disclosure on others (21)	Impacts on others (15), Security of others (3), Safety of other (2), Child grooming (1)

enough to consider a piece of personal information as sensitive if it could reveal their *identity*.

Another concern that emerged under the privacy theme was *private information*. Within this code, data items were reported to be considered more sensitive if the owners of them preferred to keep them private. Medical histories and financial status are mainly considered private and, hence sensitive by those participants. These participants also mentioned unsolicited emails, phone calls or customized advertisements as an effect of sharing information about themselves. A particular category under this privacy concern pertained to *personal life* where preferences in life, family information or relations with partners were considered sensitive by participants.

Interestingly, respondents found some publicly available information to be sensitive due to the potential use to *access more information* about the individuals. Again this was most notable when that new information was related to the health or financial status of the individuals. One poignant example in this category was the name of a pet or mother's maiden name, information commonly used for security or password questions.

Other emergent concerns included the fear of being physically traced; data items that would allow individuals to be traced were considered sensitive by a group of participants: *"People being able to find where I live or work or steal my identity.," "you can use it to track somebody, find out other information related to what you have . . . "*.

The final code related to privacy violations was the risk of third-party sharing. Some participants considered personal information sensitive when they thought it might be shared with other groups and become more widely available than expected. This concern around third-party sharing is increasingly in line with the studies that argue that third-party access leads to privacy concerns (e.g., Pang et al., 2020).

4.1.2. Two Main Contexts of Sensitive Personal Information: Health and Finance

In addition to the themes that led participants to consider certain information as more sensitive, our analysis also identified two primary contexts that heavily dominated the responses; health and finance. Hence, it is possible to report a consensus on the sensitivity of the health and finance-related information. Participants noted that these data items were expected to be given a higher standard of protection by the systems that process them. Some responses exhibited concerns regarding health information being sold or passed to insurance companies or other bodies interested in this information. Conversely, some others worried about the impact of disclosing their health status on their financial creditworthiness or career. Some participants also found health-related information inherently very private and thus sensitive, without giving any consequence as a reason.

Finance is a significantly more common response to our question when compared with health data. Several participants provided finance-related personal information as an example of sensitive information. In addition, several other data items, outside of a finance context, were considered sensitive by participants due to their impact on participants' financial status. Even though financial loss dominates the responses, some other factors such as impacts on career and financial confidentiality also led participants to find information more sensitive.

4.1.3. Financial Problems

As discussed previously, financial concerns dominated the responses. Consequences under this theme center around *financial loss*, *financial exposure*, *risk of fraud* and *negative impacts on career*. The risk of fraud appeared to be the largest concern as many participants reported information to be more sensitive if it could enable fraudulent activities. More specific responses were given by some participants where *financial loss* was explicitly given as a concern while evaluating the sensitivity level of information. *Financial exposure*, which could be considered an overlapping area between the themes *Privacy* and *Financial problems*, was another code that emerged in the responses. Finally, when evaluating the sensitivity level, several participants reflected on the impacts on their career of disclosing financial information. Political and religious affiliations, and medical histories, were popular examples given as sensitive information that participants believed could compromise their careers or aspirations.

4.1.4. Concerns Regarding the Reactions of People

Another concern of participants observed was the interpersonal *reactions* between the individual sharing the information and the individual to whom the information was disclosed. Under this theme, the most common reaction was *embarrassment* with participants reporting that information that they found embarrassing to disclose was considered sensitive.

Medical records or being a member of protected characteristics were given as examples of sensitive information

since they were considered embarrassing for themselves or their families. Similarly, *discrimination* was another code that emerged under this category. A group of participants reported a data item to be sensitive if they believed it would invoke the prejudice or bias of others. Religious or political affiliation, sexual orientation, race, disability or genetic defects and health information were examples given as sensitive due to this concern. Disclosure of personal health information has been known to result in discrimination by employers and insurance agencies if they gain access to such information (Rindfleisch, 1997).

Participants also reported finding information sensitive if it may cause them to be judged by others. In addition to *judgement*, *reputational harm* was another factor that led participants to consider a data item sensitive. We also identified *cultural conditioning*, which some participants highlighted as "taboo" subjects within society and considered items related to those taboos more sensitive (e.g., sex life, political leanings) purely because of this societal/cultural conditioning.

4.1.5. Consequences of Disclosure on the Individual

A majority of responses under this theme exhibited answers where participants defined sensitive information as the information that could be *misused/used against them* or cause them *harm*. Some participants provided more specific answers and negative effects on *mental health* and *personal safety* or feelings such as *harassment* and *fear*.

Personal security was one of the most popular responses with participants linking sensitivity to a resulting security risk. It was not possible to differentiate in the majority of the responses if the given concern was about the individuals' physical security or digital security (e.g., "I have concerns about security," "Things which might compromise my security"). However, some responses implicitly covered it where participants gave "home address" or "bank account number" as examples. *Risk of crime* is another code in this category. Participants were aware that some personal details could be used fraudulently and considered those sensitive. It is worthy of note here that almost all the concerns given in this category were in a financial context.

There were very few responses where participants shared their concerns regarding *legal issues*. Those participants reported perceiving information as sensitive if used legally against them (e.g., "official bodies can use it to deny services."). On the other hand, one participant explicitly reported considering the *costs and benefits* of disclosing information into account while evaluating its sensitivity.

4.1.6. Nature of the Information

Some participants reported data as more sensitive due to its very nature. For example, characteristics can be given as *intimacy of data* which are generally exemplified with sexual life or other information related to personal life. Participants found these data items sensitive due to their intimate nature. Another characteristic reported was the *value of the data*, which determines to what extent others can use it as it is disclosed. For instance, passwords or passport numbers were seen as more sensitive than social media data since they are perceived as having a higher impact if misused. The *relevance* is another code that emerged which defines the relevance of the information request in the given scenario. Fairness of the request was also given as a pertinent factor: *"There are certain details I would not wish to share as I do not feel they are of relevance to the data handler."*

A small group of participants considered data items that are costly to change (e.g., home address) more sensitive than items where the cost is lower (e.g., email address). Another response, albeit relatively rare, was when the data item was related to a particular *group* identity. For example, information about minors or vulnerable groups were considered sensitive. Existing research reported that a particular data item might only be sensitive where the individual belongs to a group that often faces discrimination (Rumbold and Pierscionek, 2018). For example, gender at birth is likely to be less sensitive for those who are cisgender compared to those who are transgender.

Some participants also considered the *public availability of information* while evaluating the sensitivity of it and considered that data items that were already publicly known were less sensitive.

4.1.7. Interaction Means

Disregarding the content of the information, some participants reported another essential factor; *the person/system that the information is shared with*. We identified several participants for whom the sensitivity of information is related to the receiver of the information. For some participants, it was explicitly a matter of *trust*, a data item as more sensitive if they did not trust the person or the system to whom they are disclosing it.

4.1.8. Consequences of Disclosure on Others

In addition to the previous concerns associated with the personal consequences, several responses showed a more altruistic concern. They reported considering *Consequences of disclosure on others* while evaluating the sensitivity of data items. They expressed their concerns regarding the *security and safety of their families or beloved ones*. They perceived information sensitive that could cause a risk to the security and safety of others. We have combined the generic concerns under the code *Impact on others* where participants provided their concerns without explicitly defining the impact. Most of these respondents stated that they would not share any information that would put people they know in trouble and consider these data items sensitive.

4.2. RO2: Sensitivity Rankings of Various Data Items

Beyond the factors that are taken into account while assessing the sensitivity of the information, we asked participants to rate 40 data items on a 6-point symmetric Likert scale from "not sensitive at all" (1) to "very sensitive" (6).

The participants' ratings for each data item are displayed in **Figure 2**, the data items are ordered by the average rating. Our results showed that passwords represented the most sensitive data type for UK citizens, with 92% of participants giving it the highest rating, followed by *bank account credentials* and *credit card number*, with 87 and 83% of respondents giving it the highest rating. The following items are formally identifiable information, namely national ID number and passport number, which match

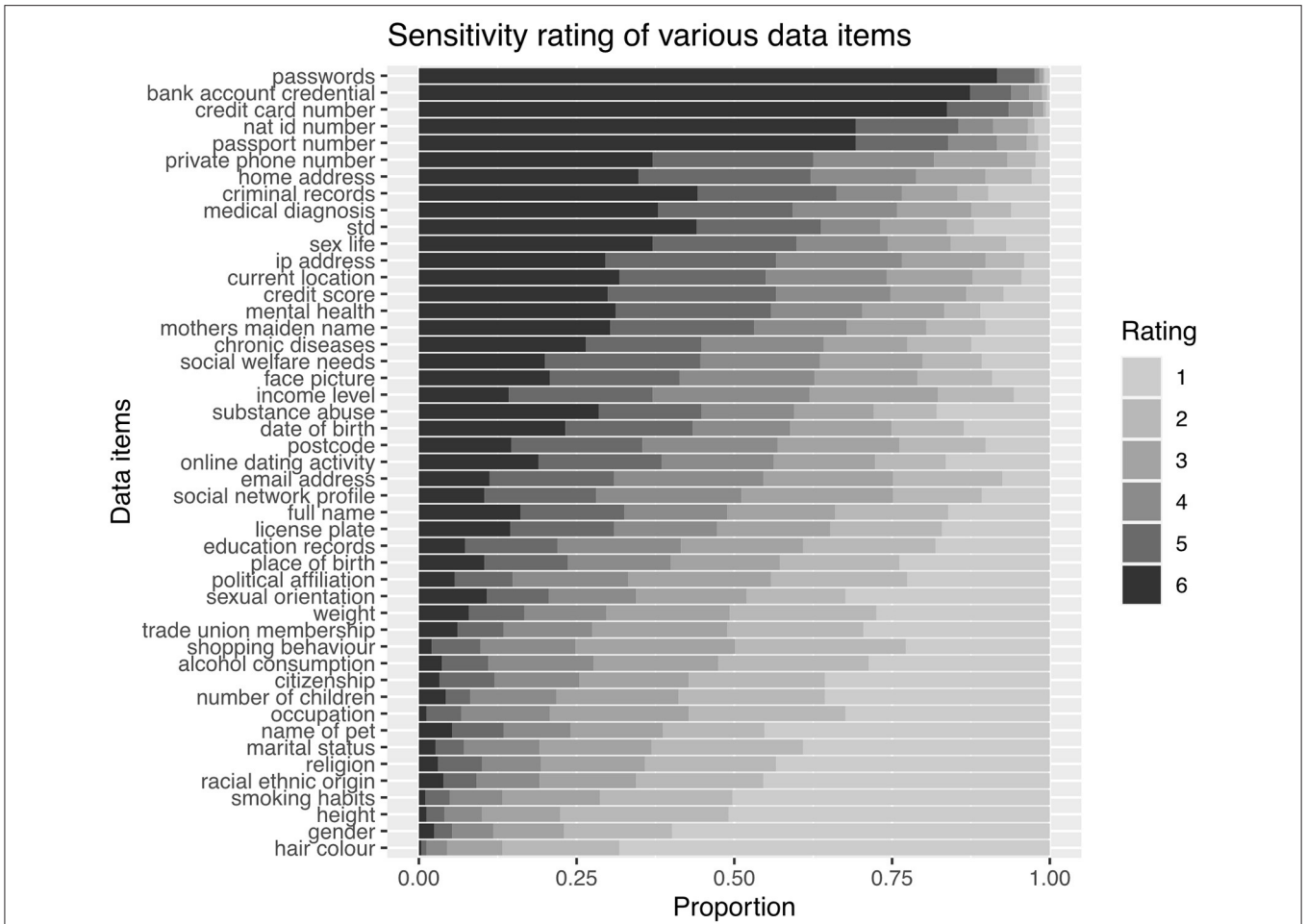


FIGURE 2 | Sensitivity ratings of data items.

the concerns given regarding identity from the first part of the questionnaire. The least sensitive items were hair color, gender and height, which are typically observable human characteristics.

4.3. RO3: Influence of User Factors

In order to examine the influence of user factors (age group, gender, education) on the perception of sensitivity, we built a proportional odds logistic regression model for each data type. We identified those data items which demonstrated a sensitivity that had a statistically significant effect (using a p-value less than 0.05) from one of these factors.

The gender of the respondents was a modulating factor on the perception of the sensitivity of an *income level*, with female respondents typically considering the sensitivity higher than male participants, (see **Figure 3**). This was also true for *IP address*, *criminal records*, *weight* and *sexually transmitted disease*. Conversely, male participants considered *smoking habits* and the *number of children* to be more sensitive than female participants.

The data items on which education has significant impact are *current location*, *political affiliation* and *sex life*. The level of education led to the sensitivity being perceived as higher for *political affiliation*. Education also modulated the perceived sensitivity of the *current location* with those who left education before achieving a post-16 qualification identifying a significantly lower sensitivity, also seen in the sensitivity of the *sex life* data item. Note, this analysis controlled for the age variable, so this is not an artifact from age measures.

The respondents' age was also observed to have significant effects on perceived sensitivity. The *Credit score* was considered significantly less sensitive by the majority of the participants aged between 18 and 24. This age group also tends to consider *date of birth*, *email address* and *mother's maiden name* less sensitive when compared to other older groups. Looking across these final three data items with factors that have a relationship with age, there tends to be an increase in sensitivity with age until the 45–54 age group before decreasing in the 55 plus age group.

4.4. RO4: Exploring Cultural Differences via Cluster Analysis

We conducted a cluster analysis on the sensitivity of the data items as done by Markos et al. (2017) and Schomakers et al. (2019). However, we used hierarchical clustering in order to gain

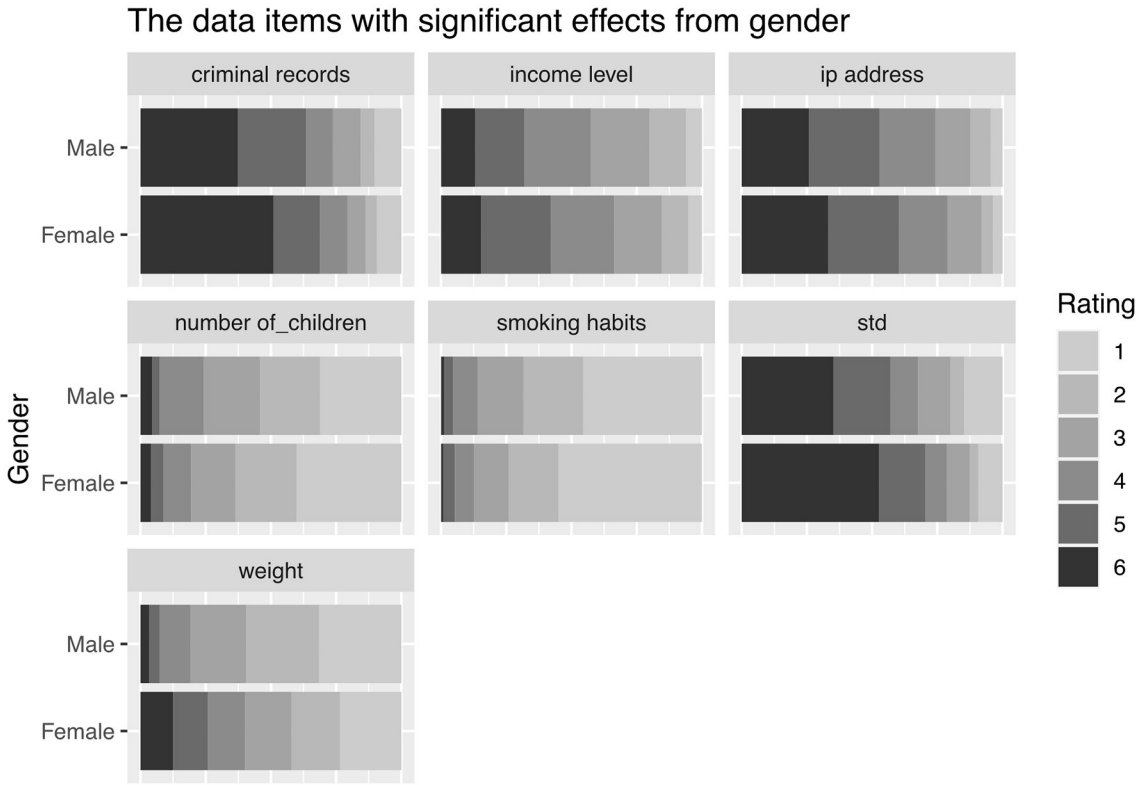


FIGURE 3 | The data items with significant effects from gender.

a high-fidelity understanding of the relationship between data items; the result is shown in **Figure 4**. Using a silhouette analysis, we found four clusters to be the most appropriate for our data set. Each cluster was cross-referenced with the ranking in **Figure 2** to label the four clusters of data categories (very highly, highly, medium and low sensitive) as shown in **Table 6**. Previous work heuristically categorized data items into three groups as highly, medium and less sensitive. However, our empirical clustering result differentiated a small group of data types from the other highly sensitive data. We grouped those items under the title of "Very highly sensitive data" in our categorization.

When previous research compared international measures of data sensitivity (Schomakers et al., 2019) it was reported that there was only one difference regarding the high sensitivity data category when they compared their results with Markos et al. (2017), which largely revealed a consensus between three countries (US, Brazilian and Germany) in this category. We see similar results with data types considered highly sensitive by those countries also appeared in the same category (or in the "Very highly sensitive data" category) in our study. In our study, several additional items appeared in this category, notably *Income level*, *current location*, *private phone number*, and *home address* were considered highly sensitive. In contrast, they belonged to medium or even low sensitive data in the German, Brazilian and US data sets. In our study, the categorization for *Credit score* was the same with the Brazilian and US data set, which differs from the medium sensitivity given by German citizens.

Among the items UK citizens placed in a medium sensitive data category, five items (*mother's maiden name*, *license plate number*, *email address*, *social network profile*, *face picture* and *post code*) were in the low sensitivity data types for German citizens. However, *mother's maiden name*, *social network profile* and *face picture* were medium sensitive not only for UK citizen but also for US and Brazilian citizens. The vehicle license plate number appeared in the medium category in our results yet was considered highly sensitive by US and Brazilian citizens and low by German citizens. The categorization of the postcode and email address was identical across all nationalities.

It is possible to report an international consensus on the low sensitive data items. Almost all data types in this category in our study were ranked into the same category as previous studies. The only difference is *sexual orientation* which was given a medium sensitive by German citizens.

4.5. RO5: Impact of the Context on Information Disclosure

The initial analysis focusing on the relationship between context and comfort in disclosing information is largely in agreement with the literature. The size of the effects is the largest seen in the study. The analysis of the data items across all scenarios is shown in **Figure 5**. In this figure, a positive model effect

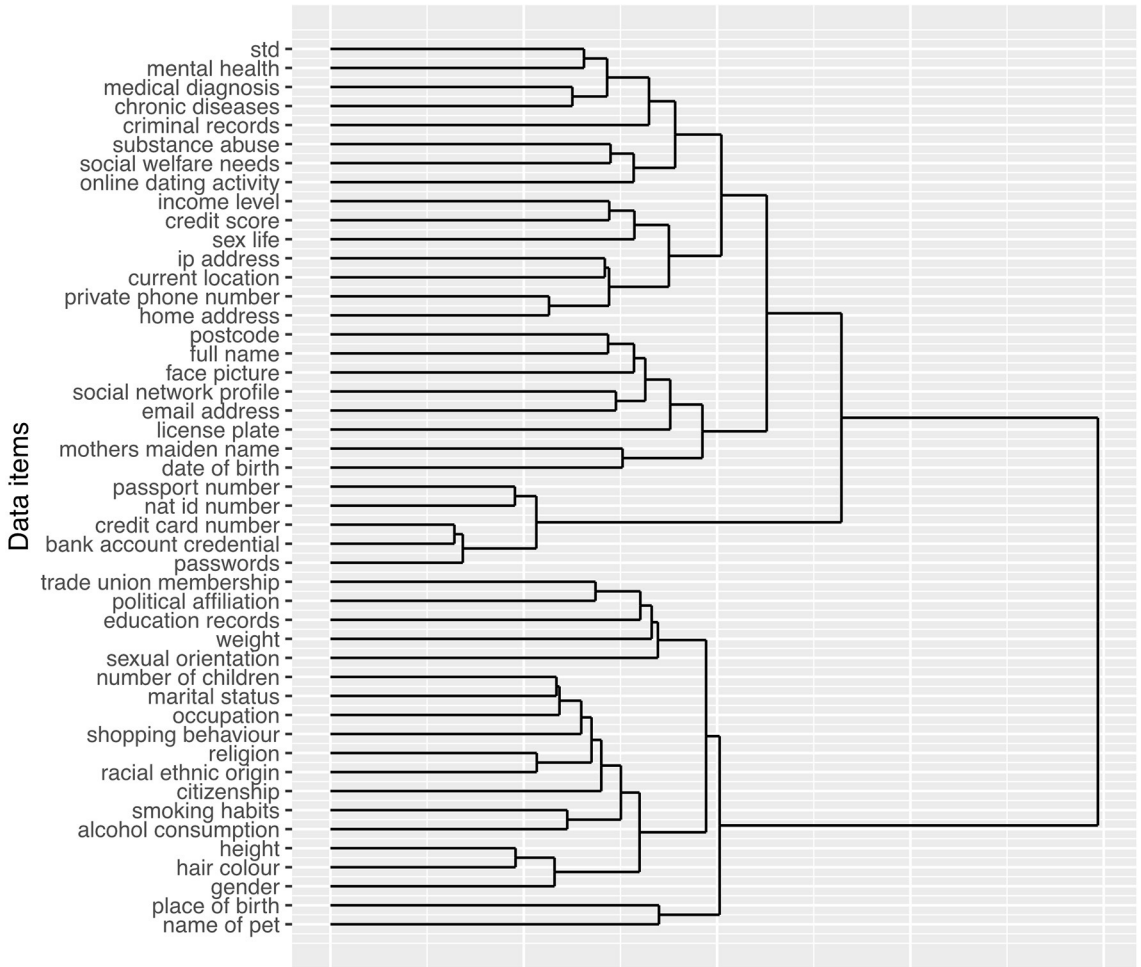


FIGURE 4 | GCSEs are the qualifications taken in years 10 and 11 of secondary school in the UK (typically at the age of 14–16). A-levels are a subject-based qualification offered by the educational bodies in the UK to students completing secondary or pre-university education.

shows participants being more comfortable disclosing in a health context and a negative model effect showing participants being more comfortable disclosing in a finance context.

There is a clear separation between the information domain and the disclosure domain, with all finance information showing negative model effects (more comfort in disclosing within a finance domain); however, there are noteworthy data items with smaller effects. There was a statistically significant effect on ethnic origin and religion where participants were more comfortable disclosing this within a health context than in the finance context. Also of interest is the small but significant effect on disclosing a criminal record; participants were more comfortable disclosing in the finance domain. However, this could be related to regulations surrounding the requirement for accurate disclosure of information in such cases.

Following a similar analysis to the previous section, we considered the pairwise comparison between scenarios S1 and S2, S3 and S4, and S5 and S6 (from **Table 2**). This results in the measures of the effect of the domain in three different scenarios: disclosing anonymously to a chatbot, disclosing anonymously to a human, and disclosing non-anonymously to a human. The effect of domain across the data items is shown in **Figure 6**.

This scenario-centered analysis clearly shows the strength of the domain effect. The domain effect is common throughout all interaction means and degrees of anonymity. An analysis of the models shows no data items where this domain effect is modulated by interaction or anonymity, and there seems to be no mechanism to significantly override or reduce this effect.

4.6. RO6: Impact of Interaction Means on Information Disclosure

The sixth research objective focused on the interaction means that elicited the disclosure; the model coefficients from the analysis of each data item are shown in **Figure 7**. Nearly two-thirds of the data items show a positive model coefficient (at a 95% confidence), indicating participants were more

TABLE 6 | Clusters of data items based on sensitivity.

Very highly sensitive data	Highly sensitive data	Medium sensitive data	Low sensitive data
Passwords	Private phone number	Date of birth	Name of pet
Bank account credential	Home address	Mother's maiden name	Place of birth
Credit card number	Current location	License plate number	Gender
National id number	IP address	Email address	Hair color
Passport number	Sex life	Social network profile	Height
	Credit score	Face picture	Alcohol consumption
	Income level	Full name	Smoking habits
	Online dating activity	Post code	Citizenship
	Social welfare needs		Racial ethnic origin
	Substance abuse		Religion
	Criminal records		Shopping behavior
	Chronic diseases		Occupation
	Medical diagnosis		Marital status
	Mental health		Number of children
	Sexually transmitted disease		Sexual orientation
			Weight
			Education records
			Political affiliation
			Trade union membership

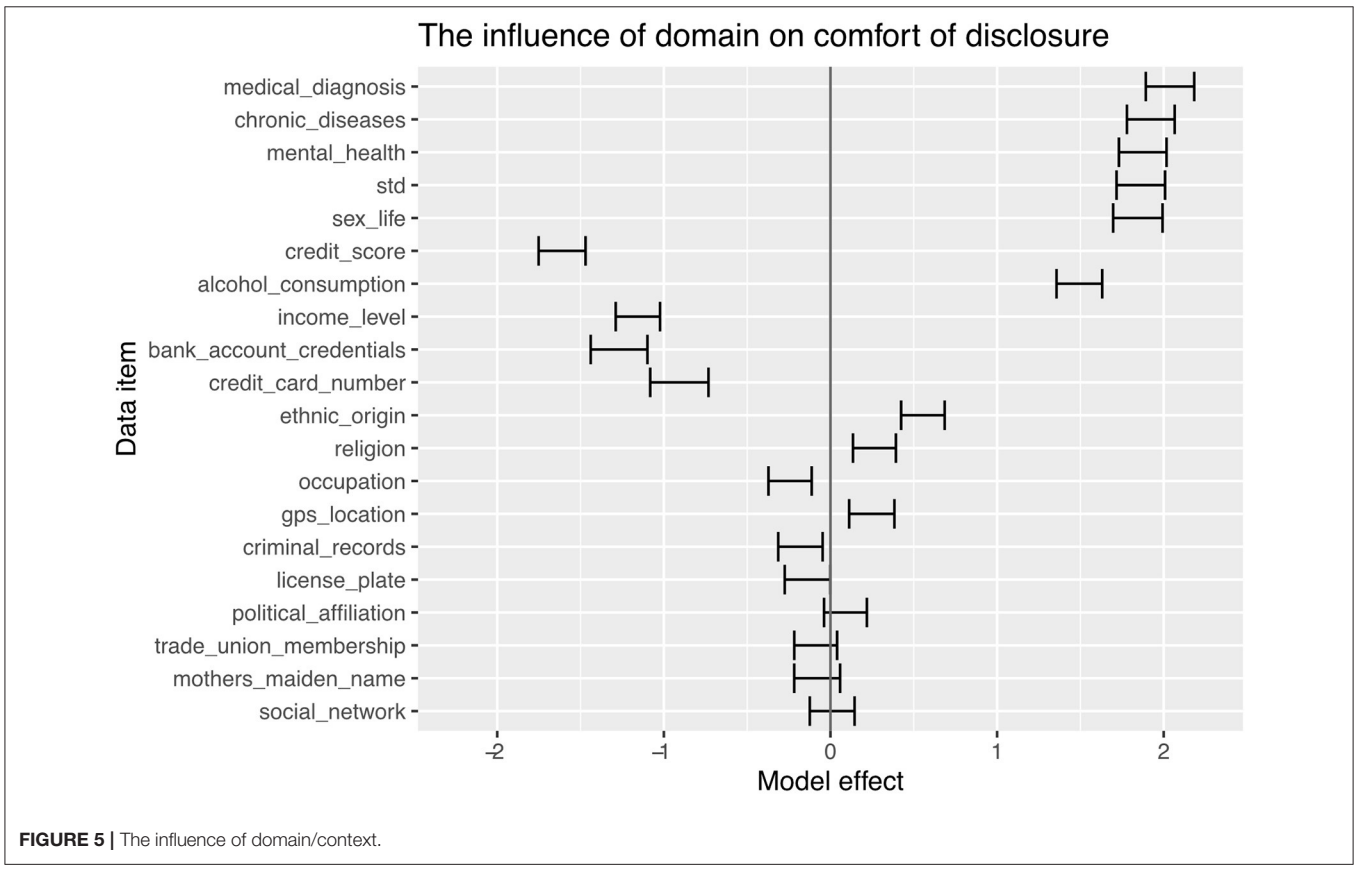


FIGURE 5 | The influence of domain/context.

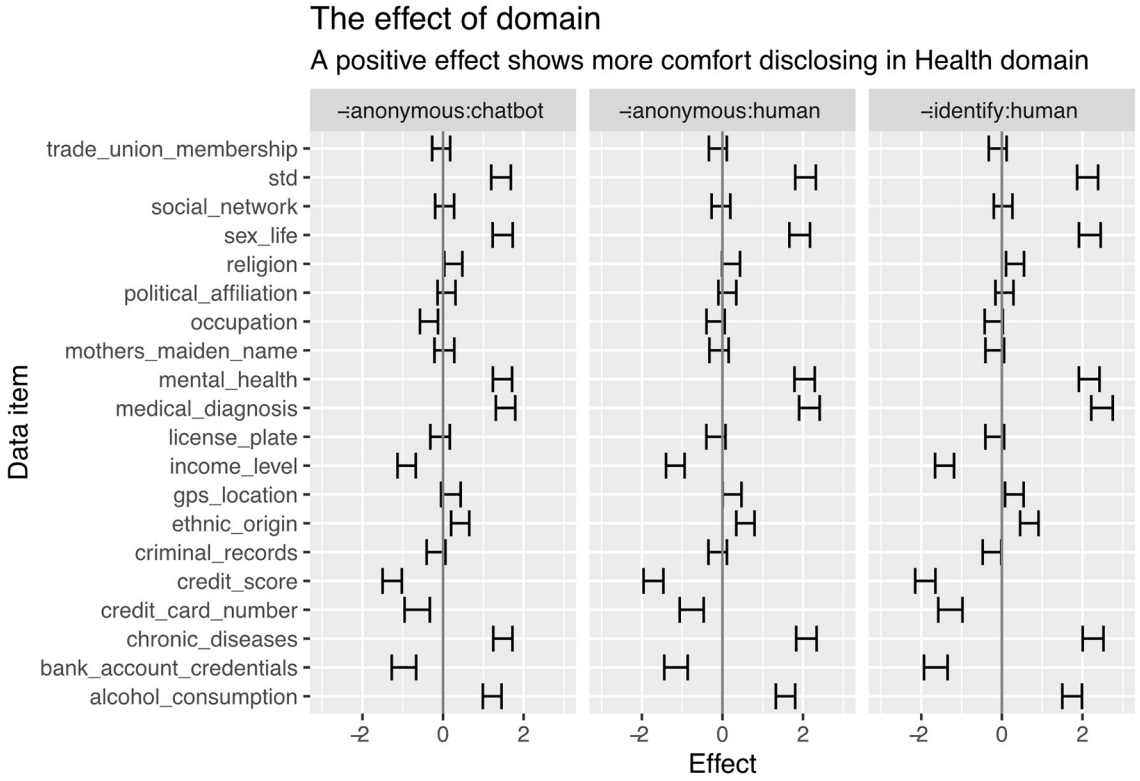


FIGURE 6 | The influence of context across different scenarios of information disclosure.

comfortable disclosing to a human than a chatbot. There were no data items that participants preferred to disclose to machines rather than humans. There was no effect from any of the biographic measures (such as age, gender and education).

Using the same modeling approach, we compared the impacts of interaction while disclosing personal information in health and finance anonymously. To achieve this, we paired the data from scenarios S1 and S5 and scenarios S2 and S6 (shown in **Table 2**). We then created a multinomial logistic regression to predict the perceptions of the sensitivity of a data item as a function of the interaction means (chatbot or human). The model coefficients are shown in **Figure 8**, with a positive effect being related to more comfort in disclosing to a human than to a chatbot (the error bounds represent the 95% confidence limit).

From these results, we observe that participants felt more comfortable disclosing sensitive information to humans, particularly in the health context. Sexually transmitted diseases, sex life, mental health, medical diagnosis or chronic diseases are data items that were preferred to be disclosed to a human by our participants. However, we can interpret this as preferring to talk to real people rather than chatbots when they need empathy and rapport in the dyadic.

Within the finance domain, only the credit score and income level data items showed a significant effect (with a 95% confidence) with interaction means. We can argue that using a chatbot will have a more negligible effect on the disclosures we would expect to be made within the finance domain.

4.7. RO7: Impact of Anonymity on Information Disclosure

This analysis considered the effect of anonymity on the disclosure of sensitive information. The logistic regression model coefficients are shown in **Figure 9**. A positive model effect related to greater comfort in disclosing when non-anonymous (i.e., the individual is identified) and a negative model coefficient demonstrates greater comfort in disclosing when the participant was anonymous.

The effect of anonymity is much smaller than other factors in this analysis. However, it does provide statistically significant effects for several data items, most notably sex life and sexually transmitted disease. Interestingly, this also includes political affiliation and alcohol consumption.

Two data items that showed a positive model effect (more comfortable in disclosing when done non-anonymously) were the mother's maiden name (something intuitively related to identity) and bank account credentials.

Considering the scenario-specific evaluation, we paired scenarios S1 and S3, and S2 and S4 to identify the effect of anonymity within the two contexts when disclosing to a human. The model effect is shown in **Figure 10** with a positive model coefficient being related to more comfort in disclosing

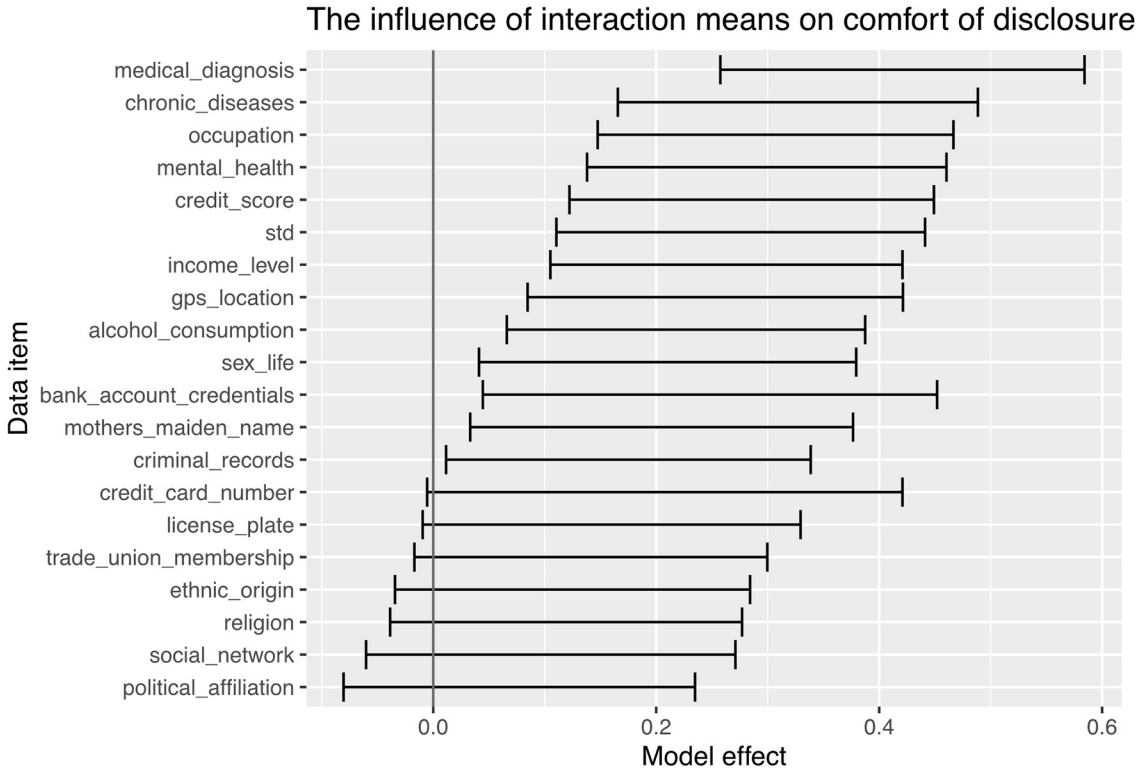


FIGURE 7 | The influence of interacting with a human or chatbot on comfort of disclosure.

when identified a negative effect coming from more comfort in disclosing when anonymous.

From these results, we can see a small effect from anonymity across the two scenarios. Within the health domain, there is a small effect associated with the sex life data item, but broadly there are very few significant effects associated with this domain. When considering the finance domain in **Figure 10** there are minor effects associated with some data items noted in the previous broader analysis. There is also a small negative effect associated with the disclosures associated with sex life in the finance domain; however, this is an out of domain disclosure whilst significant, this is likely to be an unusual disclosure.

5. DISCUSSION

In this section, we summarize and discuss our key findings for each research objective outlined previously. Furthermore, we consider the novelty of this work as compared to existing research in the field.

5.1. The Factors That Make Information Sensitive for UK Citizens (RO1)

The first research objective was to investigate the primary factors that lead British citizens to regard information as sensitive. Our findings demonstrate that there are three key general topics of note; concerns about the potential consequences of disclosure (this relates to themes *privacy*, *financial problems*, *reactions*, *consequences of disclosure on me*, *consequences of disclosure on others*), the fundamental nature of the information (themes *context*, *nature of information*), and concerns regarding the person/system the information is shared with (theme *interaction means*).

For those with privacy concerns, the main code identified was identity theft. Identity theft, the act of obtaining sensitive information about another person without their knowledge, and using this information to commit theft or fraud, is estimated to cost the UK around £190 billion every year (National Crime Agency, 2021). CIFAS, a UK-based Fraud Prevention Services, stated that in 2019, more than 364,000 cases of fraudulent conduct were recorded on their National Fraud Database with an increase of 13 per cent compared to 2018 (CIFAS, 2019). It is promising to observe the degree of awareness of this risk within the UK population; acknowledging that awareness is only the first step to prevention.

In addition, we identified several participants' decision-making was related to financial implications, with concerns regarding financial loss being one of the significant codes that emerged from the qualitative analysis. Those findings are reinforced by the items which received the highest sensitivity ratings in the quantitative phase of the study. The bank account credential, credit card number appeared in the top three most sensitive items (see **Figure 2**). They also confirm prior study which reported the possibility of harm as one of

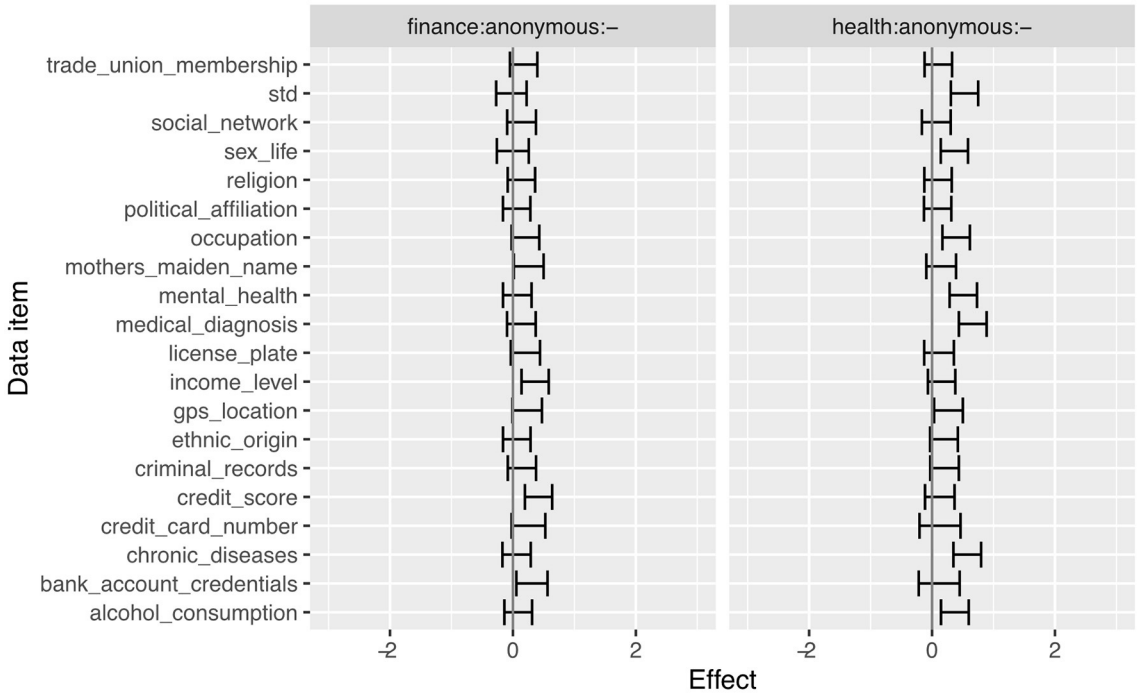


FIGURE 8 | The influence of interaction means across different scenarios.

the main factors considered when assessing sensitivity (Ohm, 2014).

Our results also uniquely highlight another concern that is generally overlooked by the privacy studies or regulations: disclosure of information belonging to others and impacts on personal information disclosure on others. Responses revealed that some participants consider information sensitive if this information belongs to others. Personal information studies in the literature are generally self-disclosure studies where the information is assumed to belong to the participant. It is also the same for the sensitivity studies where the owner of the information is assumed to be the person whose opinion or behavior is observed. Our analysis identifies concerns regarding both data belonging to others and the effect of information disclosure on others, particularly the potential harms to others. This observation indicates a societal maturity in identifying the second-order effects of disclosure.

As seen in **Figure 2**, personal data items categorized in a special category by the GDPR were not identified as being sensitive by our participants. We can identify the sensitivity of political affiliation, sexual orientation and trade-union membership as similar and not regarded as very sensitive; for example, a similar ranking was exhibited by weight and a much lower ranking than, for instance, income level or credit score. More interestingly, religion and ethnic origin were considered a very low sensitivity similar levels as marital status or occupation. Here it is worthy of note that, as mentioned before, this research aims to provide a British perspective on information sensitivity. It is well-understood that the perceived sensitivity of a particular type of data varies widely, both between societies or ethnic groups and within those groups (Rumbold and Pierscionek, 2018). The agency individuals have to protect their data, and hence the vulnerability of the individuals data affect the perceived sensitivity. Some of the data items categorized as special category by the GDPR (e.g., racial or ethnic origin or religion) may well have attracted higher sensitivity rankings if this study was constrained to minority ethnic groups rather than the general public.

5.2. Influences of User Factors on Perceived Sensitivity (RO2)

Our study also allowed us to identify variability in the perceptions of the sensitivity of data items based on the data subjects biographic information. For example, when we considered the age of the data subject, we found several interesting effects. Our findings are partially consistent with the literature that generally report that younger age groups share more information and are less concerned about information privacy, e.g., Miltgen and Peyrat-Guillard (2014) and Van den Broeck et al. (2015). It is also consistent with the literature that privacy is the most common barrier for older people to use smart technologies (Harris et al., 2022).

We can enrich those findings with fine-grained data items; for example, "credit score" was ranked less sensitive by those

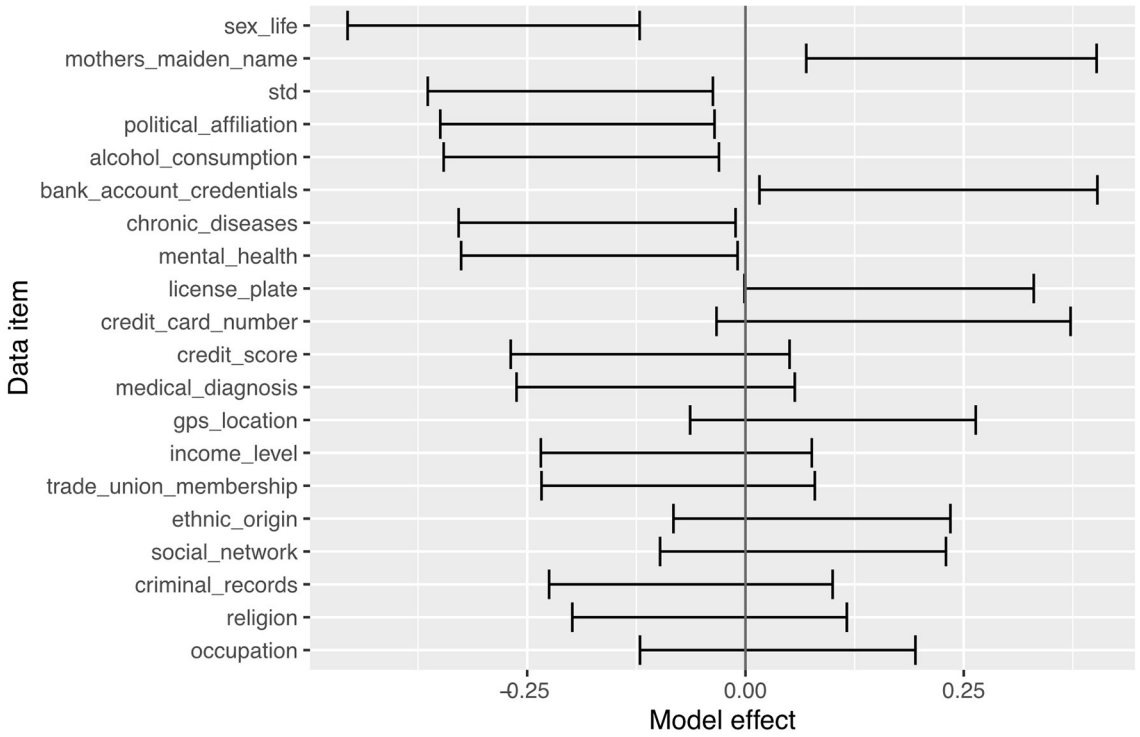


FIGURE 9 | The influence of anonymity on information disclosure.

under 25. We hypothesize that this is because this group do not normally require high credit levels (for example, purchasing a house) and hence are unlikely to be discriminated against based on that level. The same can be said of date-of-birth, which steadily becomes more sensitive during working age until retirement when it becomes less sensitive. Again there is a clear parallel with discrimination within the workplace. We believe that our detailed findings can help develop individually tailored information collection systems that recognize and respect different privacy concerns among different demographics groups.

The final two data items that show an effect with age are email address and mother's maiden name, both of which show a low sensitivity for 18–24 years with a higher level across the other age groups but with a peak in the 45–54 cohort. The reduced level of sensitivity associated with young people can be explained by the peak in the group representing Xennials or late Gen X who had an analog childhood but digital adulthood and have retained some of the understanding of the formative years of digital life. Older participants potentially have come to digital life when the internet and digital socialization norms are more formed rather than growing up alongside the transformation.

When it comes to the impact of education levels on perceived information sensitivity, we found several conflicting findings in the literature. While there are studies that claim that individuals with lower educational levels tend to be less concerned about their personal information, e.g., Rainie et al. (2013) and Blank et al. (2014), there are also those which report no differences in privacy concerns depending on education levels (Li, 2011). Our study highlights that differences in the perception of sensitivity based on education are only prevalent regarding some information types (e.g., current location, political affiliation and sex life). Within the education level, there does appear to be a breakpoint between those who achieved post-16 education, most notably in location and sex life; note this has been controlled for participant age.

The final biographic element we explored was the effect of gender on perceptions of sensitivity. Gender provided the largest number of data items that were modulated by this factor. Our study identified an apparent social stigma that female participants felt when disclosing criminal records, sexually transmitted diseases, and weight. We can also explain the higher perceived sensitivity rating of *income level* in female participants by cultural factors, which can be different in a more patriarchal society. Even though the UK is one of the countries where the lowest levels of legal discrimination are measured against women (Georgetown University's Institute for Women, Peace and Security, 2020) there is still a disconnect between the genders in terms of pay, and it naturally follows that there is a difference in the perceived sensitivity.

Our results appear to support (Knijnenburg et al., 2013) who hypothesized that information disclosure behaviors consisted of multiple related dimensions and disclosure behaviors do

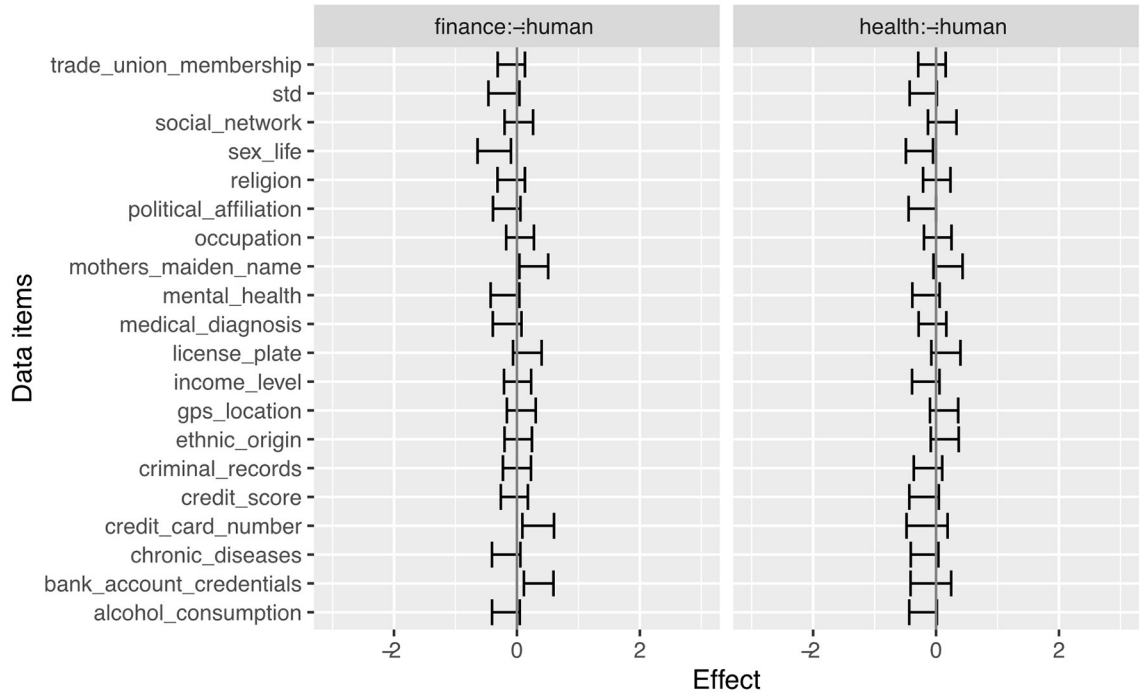


FIGURE 10 | The influence of anonymity across different scenarios of information disclosure.

not differ among groups overall, but rather in their disclosure tendencies per type of information. The results are also consistent with the results from RQ1.

5.3. UK Perspective on the Sensitivity of the Different Data Items and Identification of Cultural Differences (RO3 and RO4)

Our results confirmed the consensus on the high perceived sensitivity of the finance-related information and identifiers, which appeared in the same category as Markos et al. (2017) and Schomakers et al. (2019). When we reflect on the least sensitive items (hair color, gender, height), the common feature is that they are typically visible to the public. These appear consistent with the hypothesis from Markos et al. (2018) who predicted that public information is considered less sensitive compared to private-self information (inner states, personal history, and specific features of the self).

We conclude a degree of consensus on what constitutes sensitivity across German, US, Brazilian and UK citizens. However, respondents in our study and our rigorous empirical approach identified several "very" highly sensitive data items that formed a discrete cluster above those seen in the other studies. We also saw several elements promoted to the high-sensitivity cluster (e.g., income level, private phone number) compared to other nations, even compared to another western European country. This discontinuity shows that whilst international regulatory frameworks are undoubtedly essential to provide a degree of data protection, we must also have mechanisms to support the cultural differences within individual nations. Considering the internationalized nature of today's information society, we believe that such findings are important to consider while designing information systems that allow trans-border data flows, or for those systems designed and built in a different socio-economic environment to which they will be deployed.

5.4. Impact of the Context on Information Disclosure (RO5)

Our fifth Research Objective focused on the effect of context on the comfort of disclosing information. Our results broadly align with the literature; however, we highlight the magnitude of this effect; the strength of this effect is nearly ten times greater than any other identified in the study. **Figure 5** clearly shows that health-related information is shared with significantly more comfort in a health context. Similarly, the finance-related information is shared more comfortably in a finance context. Also interesting were the data items related to religion and ethnic origin, which exhibited significant preferences for disclosure in the medical domain. It is conceivable that ethnic origin may result in a predisposition to certain illnesses (Cooper, 2004) and justifies a disclosure in the health domain; it is unlikely that the same is true in the financial domain. The effect of context is also not mediated by the scenario and appears to be consistent

whether disclosing anonymously to a human or a chatbot or disclosing non-anonymously to a human; this is shown in **Figure 6**. These findings confirm the impact of relevance on the perceived sensitivity. From a regulatory perspective, this could be interpreted as a clear validation of the *data minimization principle* of the GDPR, which requires data collection to be adequate and limited to what is necessary.

5.5. Interaction Means and Comfort to Disclose (RO6)

Our penultimate research objective (RO6) focused on the interaction means whether the disclosure was direct to a human or through a chatbot mediated communication. In general, we found participants were more comfortable disclosing directly to a human rather than a chatbot; this was particularly the case with medical diagnosis, chronic diseases and mental health issues, shown in **Figure 7**. This preference for face-to-face human reporting has been seen in many sensitive domains, for example, within community reporting associated with violent extremism (Thomas et al., 2020). In these cases, it is very often difficult for the individual to make the disclosure. The natural interaction between humans and the perception of control is essential to support and enable these disclosures.

When this interaction means is considered in the scenario-specific conditions, we see a slightly more complicated picture. Within the health-based scenario, our participants still prefer disclosing to a human over a chatbot. Again the locus of control and the perception of engaged feedback may encourage participants to be more comfortable disclosing to a human. The other data item that showed a preference was occupation. Those findings contradict with the literature where users were reported to prefer chatbots or to respond with more disclosure intimacy to chatbots than a human (Bjaaland and Brandtzaeg, 2018; Ho et al., 2018). We can hypothesize at this point that within a healthcare setting, the perception of discussing and enriching the disclosure and providing more background as to the day-to-day tasks may drive this preference. When we consider the finance scenario, we generally see little difference between disclosure to a human or a chatbot. An indication that sensitive disclosures in this domain are less likely to be reduced through the use of conversational agents. The only data items that showed a significant effect were the credit score and income level; similarly to the occupation data item within the healthcare setting, we believe that this is a disclosure that the participant may view as requiring more enrichment or explanation. Hence, a factual disclosure with no interaction or feedback may be perceived as less desirable, leading to a perception of more comfort in disclosing to a human.

5.6. Anonymity and Comfort to Disclose (RO7)

The final research objective (RO7) focused on the effect of anonymity on the person making the disclosure. When considered abstractly, it was clear that several data items demonstrated a preference for anonymous disclosure, such as sex life and sexually transmitted diseases and alcohol consumption and political affiliation, which is inline with the previous findings (Schomakers et al., 2019). This observation would appear to match well to the qualitative results as well, which suggested that the reaction of others was an important element when judging whether items were sensitive or not.

As with the previous research objective, when this is contextualized within a real scenario, the results are more nuanced. We can see from **Figure 10** that there is no preference for anonymity within the healthcare setting—nearly all data items showed no significant difference in the comfort with being anonymous or identified. We have already demonstrated the strength of the context in the sensitivity of disclosures. We would suggest that the healthcare context and the professional reputation of the National Health Service in the UK lead to participants seeing no value in being anonymous. The only data item that showed a preference for anonymous disclosure was associated with sex life, which was only just significant at the 95% level.

When considering the finance domain, several preferences for anonymity were observed; these were mostly tied to disclosures related to health, although these effects are minor and only just significant. Hence it is difficult to draw a meaningful conclusion from this domain; however, it may hint that when disclosures are made out of domain, individuals may be more comfortable disclosing if anonymous.

6. CONCLUSION

This final section draws together our research contributions from our rigorous analytical study of this challenging problem.

6.1. Theoretical Contributions

Our study presents a detailed capture of the perspective of UK citizens regarding the sensitivity of personal information. Three main factors lead British citizens to assign higher sensitivity scores to data items; consequences of disclosure, nature of the information and the concerns regarding with whom the person/system the information is shared. Identity theft and financial loss are the main concerns of the individuals, which is consistent with the risk-based definition of sensitive personal information in regulatory documents. In addition, high sensitivity scores assigned to health and financially related information indicate that there is a consensus on what constitutes sensitivity across German, the US, Brazilian and the UK. However, British citizens regard some items as highly sensitive as compared to the other three countries. These discrepancies highlight the challenge of providing trans-national regulation and should be noted by those managing information security where data flows cross regulatory borders.

We also identified individual characteristics that modulate perceptions of sensitive data. We identified age, gender and education level as influencing the sensitivity of particular data items; these modulating characteristics mapped well to the qualitative explanations of the factors that made data items sensitive.

The context or the fairness of the request has the most significant impact on the comfort level felt while disclosing personal information. Disclosure of highly sensitive personal

information such as sex life, sexually transmitted disease or alcohol consumption was observed to be affected by anonymity. Participants reported disclosing those items with significantly more comfort when they do not have to reveal their identities.

This study has developed a systematic understanding of UK citizens' perceptions of sensitive information, showing a degree of consensus with previous studies and some unique insights. We particularly note the effect of the relevance of the disclosure and the effect of the interaction means, whether a human-mediated disclosure or a disclosure mediated by a conversational agent. In general, we highlighted the preference to disclose sensitive personal information to a human rather than a conversational agent. These findings should be considered in the design and management of information within systems that involve sensitive disclosures and hence sensitive data, particularly in the healthcare domain, where our findings are most significant.

6.2. Managerial Contributions

We contribute to the literature by investigating the impact of emerging technologies, particularly conversational agents (or chatbots), on the disclosure of personal data. Such disclose is a key security concern for both those disclosing their data and for organizations seeking to facilitate accurate, high-integrity disclosures. Despite the existence of studies that show the facilitator role of chatbots on information disclosure, no study, to our knowledge, has evaluated the perceived sensitivity of data items at granular level when they are disclosed to a chatbot. We also consequently identify the contexts where chatbots can enable individuals to disclose sensitive information more comfortably. In addition to providing general insights into how persons in the UK perceive sensitive information, our findings can contribute to the design of chatbots; most notably, defining an evidence-base to support agent use in the most appropriate usage contexts increasing the comfort of disclosing and ultimately ensuring more accurate responses. We specifically investigate two main contexts in our research; health and finance. These contexts have a regulatory demand for high levels of security and data protection, and are traditionally where chatbots are heavily adopted and sensitive personal information is frequently collected and processed (Stiefel, 2018; Ng et al., 2020). Our findings help demonstrate the relationship between the disclosed personal information and the context in which it is disclosed, ultimately uncovering the impact of usage context on disclosure of different data items. Finally, we explore the effect of anonymity, specifically identifying what personal data the UK public prefer to disclose anonymously. These observations provide novel insights for the information collection systems used in the UK by uncovering the factors that lead to perceptions of high sensitivity and hence the comfort (or discomfort) in the disclosure process.

6.3. Limitations and Future Work

While we believe our study was robust and has made several substantial contributions to the research, some limitations must be acknowledged. Firstly, our results represent self-reported sensitivity evaluations and may not reflect the lived behaviors of our participants. However, this approach allowed us to obtain and compare several sensitivity evaluations across several contexts. It also compares well with previous works in the field (e.g., Markos et al., 2017; Schomakers et al., 2019), which followed a similar methodological approach. However, we are aware that it might be possible to collect more accurate results when the participants assess their comfort levels while practicing the given scenarios.

Consequently, to validate our findings, our next step will explore the disclosure behaviors in an experimental context involving both human and chatbot mediated disclosures. Another issue faced in this study is the vagueness regarding the benefits of the disclosure and the perceived risk/trust to the interaction means. In our experimental approach, we intend to ensure a clear and consistent perception of the benefit of disclosure.

We also removed two scenarios from our 2 x 2 x 2 study; this meant that we could not fully explore all combinations of factors. However, this pragmatic decision has significantly improved the quality of the results and allowed us to draw some robust conclusions from the remaining six scenarios. Future work could consider the value in exploring all scenarios and thereby fully understanding all factors.

DATA AVAILABILITY STATEMENT

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

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Research Ethics Governance Department of University of Kent and Cranfield University Research Ethics Committee. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

RB-S, JN, and DH contributed to conception and design of the study and wrote sections of the manuscript. RB-S and JN collected the data. DH and RB-S performed the statistical analysis. RB-S wrote the first draft of the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

FUNDING

This research was conducted as a part of a UK EPSRC-funded project, A Platform for Responsive Conversational Agents to Enhance Engagement and Disclosure (PRoCEED) (Grant Nos: EP/S027297/1 and EP/S027211/1).

REFERENCES

Ackerman, M. S., Cranor, L. F., and Reagle, J. (1999). "Privacy in e-commerce: examining user scenarios and privacy preferences," in *Proceedings of the 1st ACM Conference on Electronic Commerce*. (Denver, CO: ACM), 1–8.

Aiello, G., Donvito, R., Acuti, D., Grazzini, L., Mazzoli, V., Vannucci, V., et al. (2020). Customers' willingness to disclose personal information throughout the customer purchase journey in retailing: the role of perceived warmth. *J. Retail.* 96, 490–506. doi: 10.1016/j.jretai.2020.07.001

Bansal, G., Fatemeh,Zahedi, M., and Gefen, D. (2010). The impact of personal dispositions on information sensitivity, privacy concern and trust in disclosing health information online. *Decis. Support Syst.* 49, 138–150. doi: 10.1016/j.dss.2010.01.010

Bansal, G., Zahedi, F. M., and Gefen, D. (2016). Do context and personality matter? trust and privacy concerns in disclosing private information online. *Inf. Manag.* 53, 1–21. doi: 10.1016/j.im.2015.08.001

Belen Sağlam, R., and Nurse, J. R. C. (2020). "Is your chatbot GDPR compliant? Open issues in agent design," in *Proceedings of the 2nd Conference on Conversational User Interfaces*. (Bilbao), 1–3.

Belen Sağlam, R., Nurse, J. R. C., and Hodges, D. (2022). Personal information: perceptions, types and evolution. *J. Inf. Security Appl.* 66, 103163. doi: 10.1016/j.jisa.2022.103163

Bell, S., Wood, C., and Sarkar, A. (2019). "Perceptions of chatbots in therapy," in *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*. (Glasgow), 1–6.

Bhakta, R., Savin-Baden, M., and Tombs, G. (2014). "Sharing secrets with robots?" in *EdMedia+ Innovate Learning*. (Waynesville, NC: Association for the Advancement of Computing in Education), 2295–2301.

Bjaaland, M., and Brandtzaeg, P. (2018). *Youth and News in a Digital Media Environment, Chapter Chatbots as a New User Interface for Providing Health Information to Young People*. Novi, MI: Nordicom.

Blank, G., Bolsover, G., and Dubois, E. (2014). A new privacy paradox: young people and privacy on social network sites. *Prepared Ann. Meet. Am. Sociol. Assoc.* 17, 1–35. doi: 10.2139/ssrn.2479938

Braun, V., and Clarke, V. (2006). Using thematic analysis in psychology. *Qual. Res. Psychol.* 3, 77–101. doi: 10.1191/1478088706qp063oa

Bridges Jr, C. C. (1966). Hierarchical cluster analysis. *Psychol. Rep.* 18, 851–854. doi: 10.2466/pr0.1966.18.3.851

CIFAS (2019). *Annual Report 2019*. Available online at: https://www.cifas.org.uk/about-cifas/annual-reports/annual-report-2019.

Cooper, R. S. (2004). Genetic factors in ethnic disparities in health. *Crit. Perspect. Racial Ethnic Disparities Late Life* 267, 269–309. Available online at: https://europepmc.org/books/n/nap11086/a2000af96ddd00182/?extid=20669464&src=med&fid=a2000af96ddd00196

European Parliament (2016). *Regulation (EU) (2016) 2016/679 of the European Parliament and of the Council of 27 April on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation)*. Official Journal of the European Union.

Georgetown University's Institute for Women, Peace and Security (2020). *Women Peace and Security Index 2019/20*. Available online at: https://giwps.georgetown.edu/wp-content/uploads/2019/12/WPS-Index-2019-20-Report.pdf.

Harris, M. T., Rogers, W. A., and Blocker, K. A. (2022). Older adults and smart technology: facilitators and barriers to use. *Front. Comput. Sci.* 41, 835927. doi: 10.3389/fcomp.2022.835927

Ho, A., Hancock, J., and Miner, A. S. (2018). Psychological, relational, and emotional effects of self-disclosure after conversations with a chatbot. *J. Commun.* 68, 712–733. doi: 10.1093/joc/jqy026

Ioannou, A., Tussyadiah, I., and Lu, Y. (2020). Privacy concerns and disclosure of biometric and behavioral data for travel. *Int. J. Inf. Manag.* 54, 102122. doi: 10.1016/j.ijinfomgt.2020.102122

Keith, M. J., Thompson, S. C., Hale, J., Lowry, P. B., and Greer, C. (2013). Information disclosure on mobile devices: re-examining privacy calculus with actual user behavior. *Int. J. Hum. Comput. Stud.* 71, 1163–1173. doi: 10.1016/j.ijhcs.2013.08.016

Kim, D., Park, K., Park, Y., and Ahn, J.-H. (2019). Willingness to provide personal information: perspective of privacy calculus in iot services. *Comput. Hum. Behav.* 92, 273–281. doi: 10.1016/j.chb.2018.11.022

Knijnenburg, B. P., Kobsa, A., and Jin, H. (2013). Dimensionality of information disclosure behavior. *Int. J. Hum. Comput. Stud.* 71, 1144–1162. doi: 10.1016/j.ijhcs.2013.06.003

Kolan, A., Tjoa, S., and Kieseberg, P. (2020). "Medical blockchains and privacy in Austria - technical and legal aspects," in *Proceedings of the 2020 International Conference on Software Security and Assurance* (Altoona, PA: IEEE), 1–9.

Levallois-Barth, C., and Zylberberg, H. (2017). "A purpose-based taxonomy for better governance of personal data in the internet of things era: the example of wellness data," in *Data Protection and Privacy:(In) visibilities and Infrastructures* (Cham: Springer), 139–161.

Li, Y. (2011). Empirical studies on online information privacy concerns: Literature review and an integrative framework. *Commun. Assoc. Inf. Syst.* 28, 28. doi: 10.17705/1CAIS.02828

Lozano, L. M., Garcia-Cueto, E., and Mu niz, J. (2008). Effect of the number of response categories on the reliability and validity of rating scales. *Methodology* 4, 73–79. doi: 10.1027/1614-2241.4.2.73

Malheiros, M., Preibusch, S., and Sasse, M. A. (2013). "'Fairly truthful': the impact of perceived effort, fairness, relevance, and sensitivity on personal data disclosure," in *International Conference on Trust and Trustworthy Computing* (London: Springer), 250–266.

Markos, E., Labrecque, L. I., and Milne, G. R. (2018). A new information lens: The self-concept and exchange context as a means to understand information sensitivity of anonymous and personal identifying information. *J. Interact. Mark.* 42, 46–62. doi: 10.1016/j.intmar.2018.01.004

Markos, E., Milne, G. R., and Peltier, J. W. (2017). Information sensitivity and willingness to provide continua: a comparative privacy study of the united states and brazil. *J. Public Policy Mark.* 36, 79–96. doi: 10.1509/jppm.15.159

Milne, G. R., Pettinico, G., Hajjat, F. M., and Markos, E. (2017). Information sensitivity typology: mapping the degree and type of risk consumers perceive in personal data sharing. *J. Consum. Affairs* 51, 133–161. doi: 10.1111/joca.12111

Miltgen, C. L., and Peyrat-Guillard, D. (2014). Cultural and generational influences on privacy concerns: a qualitative study in seven european countries. *Eur. J. Inf. Syst.* 23, 103–125. doi: 10.1057/ejis.2013.17

Murtagh, F., and Contreras, P. (2012). Algorithms for hierarchical clustering: an overview. *WIREs Data Min. Knowl. Disc.* 2, 86–97. doi: 10.1002/widm.53

National Crime Agency (2021). *Fraud-The Threat From Fraud*. Available online at: https://www.nationalcrimeagency.gov.uk/what-we-do/crime-threats/fraud-and-economic-crime.

Ng, M., Coopamootoo, K. P., Toreini, E., Aitken, M., Elliot, K., and van Moorsel, A. (2020). "Simulating the effects of social presence on trust, privacy concerns and usage intentions in automated bots for finance," in *2020 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW)* (Genoa: IEEE), 190–199.

Norberg, P. A., Horne, D. R., and Horne, D. A. (2007). The privacy paradox: personal information disclosure intentions versus behaviors. *J. Consum. Affairs* 41, 100–126. doi: 10.1111/j.1745-6606.2006.00070.x

Ohm, P. (2014). Sensitive information. *South Calif Law Rev.* 88, 1125–1196. Available online at:https://heinonline.org/HOL/LandingPage?handle=hein.journals/scal88&div=39&id=&page=

Pang, P. C.-I., McKay, D., Chang, S., Chen, Q., Zhang, X., and Cui, L. (2020). Privacy concerns of the australian my health record: Implications for other large-scale opt-out personal health records. *Inf. Process. Manag.* 57, 102364. doi: 10.1016/j.ipm.2020.102364

Peer, E., Brandimarte, L., Samat, S., and Acquisti, A. (2017). Beyond the turk: alternative platforms for crowdsourcing behavioral research. *J. Exp. Soc. Psychol.* 70, 153–163. doi: 10.1016/j.jesp.2017.01.006

Rainie, L., Kiesler, S., Kang, R., Madden, M., Duggan, M., Brown, S., et al. (2013). *Anonymity, Privacy, and Security Online*. Washington, DC: Pew Research Center.

Rindfleisch, T. C. (1997). Privacy, information technology, and health care. *Commun. ACM* 40, 92–100. doi: 10.1145/257874.257896

Rumbold, J. M., and Pierscionek, B. K. (2018). What are data? A categorization of the data sensitivity spectrum. *Big Data Res.* 12, 49–59. doi: 10.1016/j.bdr.2017.11.001

Schomakers, E.-M., Lidynia, C., Müllmann, D., and Ziefle, M. (2019). Internet users' perceptions of information sensitivity-insights from germany. *Int. J. Inf. Manag.* 46, 142–150. doi: 10.1016/j.ijinfomgt.2018.11.018

Schomakers, E.-M., Lidynia, C., and Ziefle, M. (2020). All of me? Users' preferences for privacy-preserving data markets and the importance of anonymity. *Electron. Markets* 30, 649–665. doi: 10.1007/s12525-020-00404-9

Stiefel, S. (2018). 'The chatbot will see you now': mental health confidentiality concerns in software therapy. *Sci. Technol. Law Rev*. 20.1 doi: 10.2139/ssrn.3166640. Available online at: https://heinonline.org/HOL/LandingPage?handle=hein.journals/cstlr20&div=12&id=&page=

Thomas, P., Grossman, M., Christmann, K., and Miah, S. (2020). Community reporting on violent extremism by "intimates": emergent findings from international evidence. *Crit. Stud. Terrorism* 13, 1–22. doi: 10.1080/17539153.2020.1791389

Treiblmaier, H., and Chong, S. (2013). "Trust and perceived risk of personal information as antecedents of online information disclosure: Results from three countries," in *Global Diffusion and Adoption of Technologies for Knowledge and Information Sharing (IGI Global)*, 41–361.

Van den Broeck, E., Poels, K., and Walrave, M. (2015). Older and wiser? Facebook use, privacy concern, and privacy protection in the life stages of emerging, young, and middle adulthood. *Soc. Media Soc*. 1, 1–11. doi: 10.1177/2056305115616149

Wadle, L. -M., Martin, N., and Ziegler, D. (2019). "Privacy and personalization: The trade-off between data disclosure and personalization benefit," in *Adjunct Publication of the 27th Conference on User Modeling, Adaptation and Personalization* (Larnaca Cyprus), 319–324.

Yu, L., Li, H., He, W., Wang, F.-K., and Jiao, S. (2020). A meta-analysis to explore privacy cognition and information disclosure of internet users. *Int. J. Inf. Manag.* 51, 102015. doi: 10.1016/j.ijinfomgt.2019.09.011

Zheng, X., Mukkamala, R. R., Vatrapu, R., and Ordieres-Mere, J. (2018). "Blockchain-based personal health data sharing system using cloud storage," in *IEEE 20th International Conference on e-Health Networking, Applications and Services (Healthcom)* (Ostrava,: IEEE), 1–6.

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

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

frontiers | Frontiers in Psychology

ORIGINAL RESEARCH
published: 06 July 2022
doi: 10.3389/fpsyg.2022.918306

I'll Do It – After One More Scroll: The Effects of Boredom Proneness, Self-Control, and Impulsivity on Online Procrastination

Cansu Sümer and Oliver B. Büttner*

Economic and Consumer Psychology, Department of Computer Science and Applied Cognitive Science, University of Duisburg-Essen, Duisburg, Germany

Procrastination is a common phenomenon. With the increasing ubiquity of new media, research has started to investigate the ways in which these technologies are used as alternatives to task engagement. This paper extends the literature by examining procrastinatory uses of social media, instant messaging, and online shopping with respect to boredom proneness, self-control, and impulsivity among German and Turkish samples. Regression analyses revealed that boredom proneness, self-control, and the perseverance facet of impulsivity are especially significant predictors of online procrastination in both samples. The results between the two studies differ in terms of impulsivity. The findings of this paper highlight the thus far understudied role of boredom proneness and various aspects of impulsivity in online procrastination, and demonstrate that social media procrastination, instant messaging procrastination, and shopping procrastination tendencies likely have distinct underlying mechanisms.

Keywords: online procrastination, social media, boredom proneness, self-control, impulsivity, online shopping

OPEN ACCESS

Edited by:
Martina Ziefle,
RWTH Aachen University, Germany

Reviewed by:
Antonio Aquino,
University of Studies G. d'Annunzio Chieti and Pescara, Italy
Corinna S. Martarelli,
Swiss Distance University Institute, Switzerland

****Correspondence:***
Cansu Sümer
cansu.suemer@uni-due.de

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

Received: *12 April 2022*
Accepted: *06 June 2022*
Published: *06 July 2022*

Citation:
Sümer C and Büttner OB (2022) I'll Do It – After One More Scroll: The Effects of Boredom Proneness, Self-Control, and Impulsivity on Online Procrastination. Front. Psychol. 13:918306. doi: 10.3389/fpsyg.2022.918306

INTRODUCTION

Imagine the following scenario: You sit down to finally write that paper. You prepare everything you are going to need. You create a new document on your computer. You are all set but you don't know where to start. You stare at the blank screen. Minutes go by. You tell yourself that you will find something to write while you are cleaning your desk, so you start organizing your workstation and think about the topic. Suddenly you get a new email from an old colleague. You wonder what they have been up to, so you check their Twitter profile. One thing leads to another, and you realize 3 hours have passed and you still haven't written a single word.

Procrastination using the Internet has gained considerable attention recently. Online procrastination is associated with lower academic performance, higher negative affect, and negative self-evaluation (Lavoie and Pychyl, 2001; Reinecke and Hofmann, 2016; Troll et al., 2021). The Internet provides an instant access to pleasurable short-term activities and enables task postponement and immediate stress relief (Lavoie and Pychyl, 2001). To date, studies have focused on procrastination using Facebook (Meier et al., 2016) and general media (television, the Internet, smartphones; Lavoie and Pychyl, 2001; Schnauber-Stockmann et al., 2018). Indeed, Facebook and instant messaging are used for postponing studying, getting away from responsibilities, and

putting tasks off (Quan-Haase and Young, 2010). Compared to their older counterparts, younger individuals use social networking sites (SNS) for procrastination more (Orchard et al., 2014).

Overall, literature indicates that online platforms are actively used as tools of procrastination. However, research is scarce regarding the procrastinatory uses of other common activities such as texting and online shopping. This is intriguing, given the reports of instant messaging applications such as WhatsApp being used regularly, delivering approximately 100 billion messages daily (Singh, 2020). Interestingly, some studies suggest that unconscientious individuals, who are more likely to procrastinate, tend to spend more time using WhatsApp (Montag et al., 2015). Although instant messaging is used frequently, no study to date has investigated whether it is indeed used for procrastinating. In a similar vein, online shopping has recently caught on, increasing by 19% in the last decade (Eurostat, 2022). Like social media, online shopping also provides an easy escape from work and everyday chores (Martínez-López et al., 2016), making it an attractive activity for procrastination. It is yet to be explored whether and how online shopping is used for procrastination as well. Therefore, the main goal of this paper was to seek an answer to how social media, instant messaging, and online shopping are used as tools of procrastination.

Both the general procrastination tendency and Internet use are influenced by proneness to get bored (Vodanovich and Rupp, 1999; Biolcati et al., 2018). However, it is unclear whether different forms of online procrastination are connected to boredom proneness. Thus, the second aim of this study was to investigate the effects of boredom proneness in addition to self-control and impulsivity as possible predictors of online procrastination. The contribution of this research is threefold. This paper is the first to distinguish between and examine different types of online procrastination (i.e., social media procrastination, instant messaging procrastination, and online shopping procrastination). Second, it is also the first to focus on the ways in which trait predictors (i.e., boredom proneness, self-control, impulsivity) contribute to these online procrastination tendencies. Finally, it strengthens its findings by examining these predictors across two culturally different samples (i.e., Germany and Turkey).

CONCEPTUAL BACKGROUND

Procrastination is the unnecessary postponing of the initiation or completion of an intended task despite the fact that the short-term prioritization of the delay will not outweigh the benefits of the long-term goals (Klingsieck, 2013). While students procrastinate more frequently than nonstudents (Svartdal et al., 2016), age is a more significant predictor of procrastination rather than one's student status (Wypych et al., 2018). Indeed, age correlates negatively with general procrastination (Beutel et al., 2016) and academic procrastination (Beswick et al., 1988).

Literature approaches procrastination as a state variable (i.e., procrastination behavior over a specific period) or as a trait variable (i.e., general tendency to procrastinate). In this paper, we focus on procrastination as a trait in various domains (i.e., general tendency to procrastinate using social media, instant messaging, and online shopping, respectively). We examine three predictors of procrastination: boredom proneness, impulsivity, and self-control. The tendency to feel boredom regardless of the situation causes one to perceive even the most common tasks as requiring effort and makes one more likely to procrastinate in general (Farmer and Sundberg, 1986; Vodanovich and Rupp, 1999; Mercer-Lynn et al., 2014). Moreover, procrastination is often conceptualized as the result of a self-control conflict between short-term desires and long-term goals. Dual-process accounts conceptualize self-control conflicts as a battle between the impulsive and the reflective system: Whether individuals give in to the temptation to procrastinate or not depends on the predominance of either the reflective capacities for self-control or the automatic, impulsive tendencies (Hofmann et al., 2009, 2017). Hence, procrastination may be the result of high impulsivity, low self-control, or both. In the following, we review the three predictors boredom proneness, impulsivity, and self-control with regards to the general procrastination tendency and different forms of media use.

Boredom

Boredom proneness is one construct that has been studied in relation to procrastination. Boredom is an aversive state where the individual is unable to engage their attention to the stimulus, is aware of this inability, and they ascribe the environment as the cause (Eastwood et al., 2012). Both attentional failures and a lack of perceived meaningfulness can lead to feelings of boredom (Westgate and Wilson, 2018). Irrespective of the situation, individuals with higher boredom proneness experience boredom more frequently, more intensely, and perceive their lives as more boring (Mercer-Lynn et al., 2014; Tam et al., 2021). They also perceive even the most typical tasks as requiring effort and tend to procrastinate more (Farmer and Sundberg, 1986; Blunt and Pychyl, 1998; Vodanovich and Rupp, 1999; Ferrari, 2000). Therefore, when people get bored during a task or have a stronger propensity to get bored in general, they are more likely to procrastinate.

It is argued that, as an adaptive state, boredom signals that one's current situation is no longer stimulating enough and thus it urges pursuing alternative activities (Bench and Lench, 2013). When individuals are bored, they frequently turn to their smartphones and social media as a pastime and to procrastinate (Martin et al., 2006; Blight et al., 2017; Alblwi et al., 2019; Koessmeier and Büttner, 2021). Similarly, people with higher boredom proneness use smartphones, SNS, and instant messaging applications more frequently (Matic et al., 2015; Wegmann et al., 2018). Shopping is also viewed as an escape from everyday life (Parsons, 2002). In fact, boredom is a strong motivation for visiting online stores and shopping impulsively, as higher boredom proneness leads to more impulse purchases (Sundstrom et al., 2019; Bozaci, 2020). To cope with boredom, individuals visit online stores and place items on their online shopping carts without any intention of buying (Kukar-Kinney and Close, 2010).

Given that both procrastination and boredom involve the urge to alleviate unpleasant states and that online services, such as social media and online stores, are frequently used to relieve boredom, it is surprising that no study has investigated whether boredom proneness is related to online procrastination. Thus, we addressed this gap by examining the ways in which boredom proneness is related to procrastinatory social media use, instant messaging, and online shopping, and expected that

> H1: Boredom proneness is positively related to all types of online procrastination.

Impulsivity

Impulsivity is another predictor of trait procrastination. While an impulse is a strong, specific, and automatically triggered inclination to approach or act on an immediate temptation or toward their short-term gratifications (Hofmann et al., 2009), impulsivity is a multifaceted construct. Specifically, Whiteside and Lynam's (2001) framework highlights four separate aspects of impulsivity. The first of these is perseverance, which describes the capacity to begin and stay focused on a task until its completion. The premeditation facet concerns the ability to consider the consequences of one's actions beforehand. Sensation seeking refers to openness to pursue new activities. Finally, the urgency facet is the tendency to act rashly when experiencing negative emotions. Apart from sensation seeking, all impulsivity factors seem to be related to procrastination. That is, while urgency positively relates to the general procrastination tendency, premeditation and perseverance are negatively related to it (Rebetez et al., 2018). Overall, a lack of perseverance or a lower capacity to remain focused on a task until its completion is the strongest predictor of procrastination (Wypych et al., 2018). In addition to this multifaceted framework of impulsivity, alternative or composite conceptualizations of this construct, such as a trait that "indicates spontaneity and a tendency to act upon whims and inclinations," are also associated with the general procrastination tendency (Steel, 2007, p. 69).

In general, higher impulsivity is associated with problematic behaviors, such as problematic uses of the smartphone, instant messaging apps, and social media (Billieux et al., 2008b; Sindermann et al., 2020). Individuals with higher impulsivity use instant messaging more automatically and for longer hours (Levine et al., 2013; Bayer et al., 2016) and more specifically, those with higher urgency send more SMSs daily (Billieux et al., 2008b). Finally, as an aspect of impulsivity, higher urgency is also the only predictor of compulsive buying (Billieux et al., 2008a). In summary, higher impulsivity and more specifically a heightened urgency is one predictor of actual and problematic media use and shopping.

Research is limited on whether impulsivity is related to online procrastination. By following Whiteside and Lynam's (2001) impulsivity framework, we aimed to examine the effects of different aspects of this trait on online procrastination. Literature suggests that lower perseverance is the strongest predictor of procrastination in general. We had no a priori hypotheses about which aspects of impulsivity would be more important for specific types of online procrastination. Therefore, we expected that

> H2: The perseverance aspect of impulsivity is negatively related to all types of online procrastination.

Self-Control

Self-control is the ability to willfully adjust behaviors when one's abstract or remote goals (e.g., getting good grades) are conflicted by more concrete or immediate desires (e.g., going on Instagram to see what is new) and to refrain from acting on the latter (Tangney et al., 2004; Fujita, 2011). In conditions where the strength of the impulsive system increases, the reflective system may fail to inhibit and override impulses, whose "activation level exceeds the critical threshold necessary for the execution of self-controlled behavior" (Hofmann et al., 2009, p. 165). This capacity to successfully deal with problematic desires that conflict with one's goals is crucial for task completion and procrastination (Sirois and Pychyl, 2013; Pychyl and Sirois, 2016). Indeed, self-control is negatively related to the general procrastination tendency (Wijaya and Tori, 2018). In short, self-control prevents one from giving in to the temptation of quitting the task in favor of more pleasant alternatives or not engaging with it at all.

The ever-present availability of media poses a challenge for media consumers' goals and task completion in everyday life (Hofmann et al., 2017), which is why self-control is one of the constructs that have been most commonly examined in relation to media use. Specifically, self-control is negatively associated with habitual Facebook checking as well as the duration of media use, including daily instant messaging, SNS, TV, and online videos (Panek, 2014; Li et al., 2016; Meier et al., 2016). Lower self-control also predicts problematic online shopping and compulsive buying (Achtziger et al., 2015; Jiang et al., 2017). Thus, difficulties with successfully handling desires for media use in favor of higher-order goals can lead to procrastination.

Similar to the negative associations between self-control and general procrastination tendency (Wijaya and Tori, 2018), research indicates that self-control is negatively related to procrastinatory uses of Facebook (Meier et al., 2016), smartphones (Schnauber-Stockmann et al., 2018; Troll et al., 2021), and general media (the Internet, TV, video games; Reinecke and Hofmann, 2016). Similarly, ego depletion, that is, the "temporary reduction in the self's capacity or willingness to engage in volitional action" (Baumeister et al., 1998, p. 1253), is positively associated with procrastinatory media use (Reinecke et al., 2014). Overall, these studies suggest that individuals with lower self-control are more likely to use media to procrastinate. Therefore, we also examined the effect of self-control on different types of online procrastination and expected that

> H3: Self-control is negatively related to all types of online procrastination.

Overview of Studies

To address these, two online studies were carried out. Study 1 used a quota sample and investigated how people use online services for procrastination in Germany. It also aimed to

TABLE 1 | Participant demographics.

		%	
		German	**Turkish**
Age	18–29	18.6	22.9
	30–39	19.9	27.3
	40–49	20.5	21.0
	50–59	19.9	20.0
	60–69	21.1	8.8
Gender	Male	50.2	30.2
	Female	49.8	67.8
	Other		2.0
Employment status	Employee	51.5	43.9
	In training	10.7	16.1
	Self-employed or part-time	5.2	9.3
	Job seeking	4.6	14.6
	Active in the household, retired or other	28.0	16.1
Marital status	Single	28.0	41.5
	Married or in a partnership	58.6	46.3
	Divorced, widowed or other	13.4	12.2

TABLE 2 | Means and standard deviations of variables.

	German	**Turkish**
	***M* (SD)**	***M* (SD)**
Age	44.66 (14.58)	40.13 (13.18)
Online hours	4.17 (2.82)	4.91 (3.09)
Social media hours	2.10 (2.27)	2.81 (2.34)
Self-control	3.36 (0.63)	3.31(0.69)
UPPS: Urgency	2.33 (0.76)	2.63 (0.75)
UPPS: Premeditation	3.68 (0.61)	3.94 (0.61)
UPPS: Perseverance	3.86 (0.64)	3.53 (0.77)
UPPS: Sensation Seeking	2.45 (0.93)	2.86 (0.96)
Trait procrastination	2.16 (0.73)	2.77 (0.90)
Boredom proneness	2.54 (1.16)	3.30 (1.47)
Social media procrastination	2.28 (0.97)	2.80 (1.09)
Instant messaging procrastination	2.16 (0.93)	2.83 (1.07)
Shopping procrastination	2.04 (0.90)	1.93 (0.90)

establish the individual differences in trait self-control, boredom proneness, and impulsivity in relation to different types of online procrastination. Because we wanted to explore the differences between the predictors of online procrastination across different countries, we carried out Study 2 using the same measures in a convenience sample from a different cultural background, namely, Turkey. We did not have a priori hypotheses regarding these differences and addressed this research question in an exploratory way.

STUDY 1

MATERIALS AND METHODS

Sample

Data was collected from 333 German participants through a commercial online access panel. As we aimed for a heterogeneous sample that reflects individuals with different backgrounds and experiences, we used quotas for gender and age (50% men, 50% women; 20% from each age group: 18–29, 30–39, 40–49, 50–59, 60–69 years). Additionally, we wanted to focus on individuals that were active users of social media and instant messaging. Thus, to be eligible, participants had to have at least one social media or instant messaging account and use social media at least a few times a month. Participants who finished the survey in less than 3 min, failed to complete it in a single session, and failed both control questions were excluded. The final dataset included 307 participants (M_{age} = 44.66 years, SD = 14.58, 49.8% female; see **Table 1** for participant demographics).

Measures

In addition to demographics information, participants reported how many hours they spent on social media and on the Internet, daily. Finally, they reported on the following scales.

Online Procrastination

To our knowledge, there are no standardized scales of online procrastination. Therefore, the four-item measure used by Reinecke et al. (2014) was adapted to measure procrastinatory social media use (e.g., "I use social media although I have planned to get something done"), instant messaging (e.g., "I use instant messaging although I have more important things to do"), and browsing of online shops (e.g., "I browse online shops while procrastinating upcoming work"), separately. The German translations were adapted from Troll et al. (2021). The items were rated on a five-point rating scale (αs = 0.95 for all three types of online procrastination).

Boredom Proneness

The eight-item Short Boredom Proneness Scale (SBPS; Farmer and Sundberg, 1986; Struk et al., 2017) was used to measure the tendency to experience boredom (e.g., "I find it hard to entertain myself"). The SBPS was translated into German by Martarelli et al. (2020). The items were rated on a seven-point rating scale (α = 0.91).

Trait Self-Control

The Brief Self-Control Scale (Tangney et al., 2004) is a widely used measure of trait self-control (e.g., "I'm good at resisting temptation"). It was adapted to German by Bertrams and Dickhäuser (2009) and included 13 items, which were rated on a five-point rating scale (α = 0.84).

Impulsivity

The UPPS Impulsive Behavior Scale was created by Whiteside and Lynam (2001) to capture the four facets of impulsivity through the subscales Urgency, Premeditation, Perseverance, and Sensation Seeking. A 20-item short version that includes the four subscales was adapted to German by Keye et al. (2009). The items were rated on a five-point scale (α = 0.80 for Urgency, 0.69 for Premeditation, 0.72 for Perseverance, 0.75 for Sensation Seeking).

Trait Procrastination

The Pure Procrastination Scale was used to measure chronic procrastination (e.g., "I'm continually saying 'I'll do it tomorrow'"; Steel, 2010). It was adapted to German by

TABLE 3 | Bivariate correlations in the German sample above the diagonal and the Turkish below the diagonal.

	1	2	3	4	5	6	7	8	9	10	11	12	13
1. Age		−0.10	−0.24**	0.24**	−0.21**	0.05	0.23**	−0.27**	−0.28**	−0.28**	−0.43**	−0.43**	−0.28**
2. Online hours	−0.36**		0.38**	−0.10	0.04	−0.01	−0.09	0.05	−0.00	0.11*	0.09	0.06	0.02
3. Social media hours	−0.22**	0.64**		−0.21**	0.24**	0.01	−0.18**	0.13*	0.14*	0.22**	0.31**	0.24**	0.24**
4. Self-control	0.50**	−0.40**	−0.27**		−0.61**	0.30**	0.67**	−0.15**	−0.62**	−0.62**	−0.53**	−0.51**	−0.45**
5. UPPS: Urgency	−0.34**	0.37**	0.25**	−0.66**		−0.27**	−0.56**	0.23**	0.45**	0.55**	0.42**	0.44**	0.43**
6. UPPS: Premeditation	−0.05	−0.03	−0.05	0.31**	−0.32**		0.36**	−0.05	−0.15**	−0.11*	−0.08	−0.08	−0.10
7. UPPS: Perseverance	0.46**	−0.32**	−0.21**	0.73**	−0.48**	0.29**		−0.14*	−0.68**	−0.65**	−0.51**	−0.45**	−0.48**
8. UPPS: Sensation Seeking	−0.28**	0.13*	0.04	−0.25**	0.21**	0.02	−0.08		0.15**	0.16**	0.23**	0.18**	0.17**
9. Trait procrastination	−0.41**	0.40**	0.27**	−0.70**	0.48**	−0.11	−0.73**	0.09		0.65**	0.68**	0.59**	0.58**
10. Boredom proneness	−0.48**	0.38**	0.25**	−0.58**	0.49**	−0.09	−0.59**	0.15*	0.65**		0.59**	0.51**	0.48**
11. Social media procrastination	−0.46**	0.51**	0.36**	−0.61**	0.44**	−0.11	−0.61**	0.12	0.71**	0.54**		0.78**	0.63**
12. Instant messaging procrastination	−0.37**	0.35**	0.22**	−0.53**	0.38**	−0.09	−0.50**	0.15*	0.54**	0.48**	0.68**		0.65**
13. Shopping procrastination	−0.28**	0.32**	0.14*	−0.42**	0.37**	−0.06	−0.34**	0.01	0.41**	0.41**	0.44**	0.37**	

UPPS, UPPS Impulsive Behavior Scale.
***p < 0.01.*
**p < 0.05.*

Svartdal et al. (2016) and included 12 items, which were rated on a five-point rating scale (α = 0.91).

RESULTS

Means and standard deviations can be found in **Table 2**. In order to determine whether or not to include age and gender as control variables in our further analyses, we carried out independent samples *t*-test and discovered gender differences in sensation seeking, with men (M = 2.60, SD = 0.92) scoring higher than women (M = 2.31, SD = 0.93), $t(305) = -2.67$, $p < 0.01$. Younger individuals used social media longer ($r = -0.248$, $p < 0.001$). The correlation between age and hours spent online was marginally significant ($r = -0.104$, $p = 0.069$). For exploratory purposes, we examined the correlations between trait procrastination and other variables. Trait procrastination had a higher correlation with social media procrastination ($r = 0.68$, $p < 0.01$) than with instant messaging procrastination ($r = 0.59$, $p < 0.01$) and shopping procrastination ($r = 0.58$, $p < 0.01$). See **Table 3** for further bivariate correlations. Further exploratory analyses showed that Facebook, Instagram, and YouTube were the most frequently visited websites that were used while procrastinating (see **Table 4**).

TABLE 4 | Percentage of social media sites and instant messaging applications that are used for procrastination.

		German	Turkish
Social media sites	Facebook	51.8	42.0
	Instagram	33.2	66.3
	Twitter	9.4	38.5
	YouTube	45.0	57.1
	Tumblr	2.0	18.5
	Snapchat	7.8	2.0
	Pinterest	1.0	13.7
	Other	14.0	18.5
	None	19.9	2.0
Instant messaging	WhatsApp	77.9	92.7
	Facebook Messenger	21.5	17.1
	Telegram	8.8	13.2
	Instagram Messenger	0.3	1.5
	Other	11.8	6.8
	None	18.9	4.4

To test the hypotheses, three hierarchical regression analyses were carried out with procrastinatory social media use, instant messaging, and visits to online shops as the dependent variable, separately. In all three analyses, boredom proneness, self-control, and the four impulsivity facets were entered as predictors in Step 1. Age and gender were entered as control variables in Step 2.

For social media procrastination, the final model showed that boredom proneness was the strongest predictor (β = 0.30, $p < 0.001$), followed by age (β = −0.24, $p < 0.001$), self-control (β = −0.20, $p = 0.001$), and perseverance (β = −0.13, $p = 0.044$). For procrastinatory instant messaging, age was the strongest predictor (β = −0.27, $p < 0.001$), followed by self-control (β = −0.22, $p = 0.001$), boredom proneness (β = 0.17, $p = 0.008$), and urgency (β = 0.12, $p = 0.033$). Finally, for procrastinatory browsing of online stores, perseverance (β = −0.22, $p = 0.003$) made the strongest contribution, followed by boredom proneness (β = 0.16, $p = 0.019$), urgency (β = 0.12, $p = 0.049$), and age (β = −0.11, $p = 0.025$). The detailed results of the regression analyses are available in **Table 5**.

In summary, boredom proneness was positively related to social media procrastination, instant messaging procrastination, and online shopping procrastination. Thus, H1 is supported. Perseverance was negatively related to social media procrastination and online shopping procrastination, but not instant messaging procrastination. Hence, H2 is partially supported. Finally, self-control was negatively related to social media procrastination, instant messaging procrastination, but not shopping procrastination. Therefore, H3 is also only partially supported.

TABLE 5 | Hierarchical regressions predicting online procrastination in the German sample.

	Procrastinatory social media				Procrastinatory instant messaging				Procrastinatory shopping			
Variable	***B***	**SE**	***β***	***p***	***B***	**SE**	***β***	***p***	***B***	**SE**	***β***	***p***
Step 1												
Constant	2.74	0.58		<0.001	2.37	0.60		<0.001	2.54	0.59		<0.001
Boredom proneness	0.29	0.05	0.34	<0.001	0.18	0.05	0.22	0.001	0.14	0.05	0.18	0.008
Self-control	−0.36	0.10	−0.23	<0.001	−0.37	0.10	−0.25	<0.001	−0.18	0.10	−0.12	0.077
UPPS: Urgency	0.00	0.07	0.00	0.98	0.15	0.07	0.12	0.046	0.15	0.07	0.12	0.050
UPPS: Sensation Seeking	0.12	0.04	0.12	0.006	0.07	0.04	0.07	0.114	0.06	0.04	0.07	0.148
UPPS: Perseverance	−0.20	0.10	−0.13	0.046	−0.11	0.10	−0.08	0.258	−0.31	0.10	−0.22	0.003
UPPS: Premeditation	0.13	0.07	0.08	0.074	0.13	0.07	0.09	0.082	0.11	0.07	0.08	0.130
R^2			0.428				0.342				0.318	
Adjusted R^2			0.416				0.333				0.304	
Step 2												
Constant	3.69	0.58		<0.001	3.40	0.59		<0.001	3.01	0.61		<0.001
Boredom proneness	0.25	0.05	0.30	<0.001	0.14	0.05	0.17	0.008	0.12	0.05	0.16	0.019
Self-control	−0.32	0.09	−0.20	0.001	−0.32	0.10	−0.22	0.001	−0.15	0.10	−0.11	0.123
UPPS: Urgency	0.00	0.07	0.00	0.970	0.15	0.07	0.12	0.033	0.15	0.07	0.12	0.049
UPPS: Sensation Seeking	0.08	0.04	0.07	0.088	0.02	0.04	0.02	0.617	0.05	0.04	0.05	0.299
UPPS: Perseverance	−0.19	0.09	−0.13	0.044	−0.11	0.10	−0.07	0.261	−0.31	0.10	−0.22	0.003
UPPS: Premeditation	0.13	0.07	0.08	0.069	0.13	0.07	0.08	0.075	0.11	0.07	0.08	0.128
Age	−0.01	0.00	−0.24	<0.001	−0.01	0.00	−0.27	<0.001	−0.00	0.00	−0.11	0.025
Gender	−0.11	0.08	−0.05	0.186	−0.12	0.08	−0.06	0.145	−0.09	0.08	−0.05	0.297
R^2			0.484				0.41				0.333	
Adjusted R^2			0.470				0.40				0.315	

DISCUSSION

Study 1 examined the effects of several personality traits on online procrastination in a German sample. The positive correlations between general procrastination tendency and different types of online procrastination suggest that chronic procrastinators also use social media, instant messaging, and online shopping for procrastination. Regression analyses showed that younger and more boredom prone individuals used social media, instant messaging, and online shopping more frequently to procrastinate. Moreover, individuals with lower self-control used social media and instant messaging more to procrastinate. Finally, urgency predicted procrastinatory instant messaging and online shopping, whereas perseverance predicted procrastinatory social media use and online shopping. In Study 2, we investigated the effects of these predictors in another country.

STUDY 2

MATERIALS AND METHODS

Sample

In total, 217 Turkish adults participated. Participants were reached through snowballing and word-of-mouth. Again, participants had to have at least one social media or instant messaging account and use social media at least a few times a month. The final dataset included 205 participants (M_{age} = 40.13, SD = 13.18, 67.8% female).

Measures

We used the same measures as in Study 1 in Turkish versions. For online procrastination, the four-item measure from Reinecke et al. (2014) was used again to measure procrastination with social media, instant messaging, and online shopping, separately. The items were translated by the first author and reviewed by an English-Turkish translator (αs between 0.94 and 0.96). For boredom proneness, we used the Turkish version of the short SBPS (Güner et al., 2021; α = 0.90). For trait self-control, BSCS was used in Turkish (Nebioglu et al., 2012; α = 0.85). The UPPS Impulsive Behavior Scale was used in Turkish (Yargıç et al., 2011; α = 0.71 for Urgency, 0.71 for Premeditation, 0.77 for Perseverance, 0.78 for Sensation Seeking). For trait procrastination, we used the Turkish version of the PPS (Balkis and Duru, 2019; α = 0.92).

RESULTS

Independent samples *t*-tests indicated significant gender differences in sensation seeking, with men (M = 3.47, SD = .78) scoring higher than women (M = 2.59, SD = 0.93), $t(138) = -6.96$, $p < 0.001$. On self-control, women (M = 3.41, SD = 0.68) scored higher than men (M = 3.12, SD = 0.68), $t(199) = 2.74$, $p < 0.01$. Younger individuals spent longer using both social media

TABLE 6 | Hierarchical regressions predicting online procrastination in the Turkish sample.

	Procrastinatory social media				Procrastinatory instant messaging				Procrastinatory shopping			
Variable	***B***	**SE**	***β***	***p***	***B***	**SE**	***β***	***p***	***B***	**SE**	***β***	***p***
Step 1												
Constant	4.51	0.746		<0.001	4.08	0.818		<0.001	1.94	0.739		0.009
Boredom proneness	0.13	0.05	0.17	0.011	0.14	0.05	0.19	0.013	0.13	0.05	0.21	0.008
Self-control	−0.43	0.14	−0.27	0.003	−0.45	0.16	−0.29	0.005	−0.36	0.14	−0.27	0.014
UPPS: Urgency	0.08	0.10	0.05	0.428	0.02	0.11	0.02	0.803	0.17	0.10	0.14	0.083
UPPS: Sensation Seeking	−0.01	0.06	−0.01	0.769	0.03	0.06	0.03	0.575	−0.11	0.06	−0.12	0.065
UPPS: Perseverance	−0.44	0.11	−0.31	<0.001	−0.26	0.12	−0.18	0.042	0.03	0.11	0.02	0.781
UPPS: Premeditation	0.17	0.10	0.10	0.080	0.13	0.11	0.07	0.220	0.12	0.10	0.08	0.224
R^2			0.471				0.350				0.248	
Adjusted R^2			0.455				0.330				0.225	
Step 2												
Constant	4.91	0.798		<0.001	4.18	0.882		<0.001	2.45	0.78		0.002
Boredom proneness	0.11	0.05	0.15	0.029	0.13	0.05	0.18	0.019	0.11	0.05	0.18	0.030
Self-control	−0.37	0.14	−0.24	0.012	−0.43	0.16	−0.27	0.009	−0.37	0.14	−0.28	0.011
UPPS: Urgency	0.07	0.10	0.05	0.441	0.02	0.11	0.02	0.807	0.17	0.10	0.14	0.092
UPPS: Sensation Seeking	−0.04	0.06	−0.04	0.461	0.02	0.07	0.02	0.755	−0.07	0.06	−0.07	0.289
UPPS: Perseverance	−0.40	0.11	−0.28	0.001	−0.24	0.13	−0.17	0.060	0.02	0.11	0.01	0.865
UPPS: Premeditation	0.12	0.10	0.07	0.223	0.11	0.11	0.06	0.310	0.13	0.10	0.08	0.201
Age	−0.01	0.00	−0.12	0.054	−0.00	0.00	−0.04	0.560	−0.00	0.00	−0.04	0.547
Gender	0.03	0.12	0.01	0.755	0.03	0.13	0.01	0.782	−0.26	0.11	−0.15	0.026
R^2			0.481				0.351				0.268	
Adjusted R^2			0.460				0.325				0.238	

($r = -0.224$, $p = 0.000$) and the Internet ($r = -0.369$, $p < 0.000$). Trait procrastination had a higher correlation with social media procrastination ($r = 0.71$, $p < 0.01$) than with instant messaging ($r = 0.54$, $p < 0.01$) and shopping ($r = 0.41$, $p < 0.01$). See **Table 3** for further bivariate correlations.

The same hierarchical analyses were carried out as in Study 1. Specifically, three hierarchical regression analyses were carried out with procrastinatory social media use, instant messaging, and visits to online shops as the dependent variable, separately. In all analyses, boredom proneness, self-control, and the four impulsivity facets were entered as predictors in Step 1. Age and gender were entered as control variables in Step 2.

For social media procrastination, in the final model, perseverance was the strongest predictor ($\beta = -0.28$, $p = 0.001$), followed by self-control ($\beta = -0.24$, $p = 0.012$), and boredom proneness ($\beta = 0.15$, $p = 0.029$). For procrastinatory instant messaging, self-control ($\beta = -0.27$, $p = 0.009$) and boredom proneness ($\beta = 0.18$, $p = 0.019$) were the only significant predictors. For procrastinatory shopping, self-control ($\beta = -0.28$, $p = 0.011$) made the strongest contribution, followed by boredom proneness ($\beta = 0.18$, $p = 0.030$), and gender ($\beta = -0.15$, $p = 0.026$). The detailed results of the regression analyses are available in **Table 6**.

In summary, boredom proneness was positively related to all three types of online procrastination. Thus, H1 is supported. Perseverance was negatively related to social media procrastination but not instant messaging procrastination or shopping procrastination. Hence, H2 is only partially supported. Finally, self-control was negatively related to all types of online procrastination. Therefore, H3 is supported.

DISCUSSION

Study 2 aimed to explore the effects of the same predictors in Study 1 in a different sample. As earlier, higher boredom proneness and lower self-control and perseverance predicted social media procrastination. For instant messaging procrastination, higher boredom and lower self-control increased this tendency. Unlike Study 1, in which urgency was a positive predictor, no impulsivity facet predicted this procrastination tendency. Finally, in Study 1, boredom proneness, urgency, and perseverance had predicted shopping procrastination. In Study 2, boredom proneness and self-control were the only significant predictors of this tendency.

GENERAL DISCUSSION

The aim of this paper was to extend the online procrastination literature by investigating the effects of self-control, boredom proneness, and impulsivity. We focused on procrastinatory social media use, instant messaging, and browsing of online stores in German and Turkish samples. Our findings demonstrate that higher boredom proneness, lower self-control, and lower perseverance are especially predictive of different types of

online procrastination tendencies across both samples. To our knowledge, this study is the first attempt to understand the role of boredom proneness in online procrastination. We found that a higher propensity to get bored leads to more frequent procrastination with social media, instant messaging apps, and online shops. While some trait variables (i.e., boredom proneness and self-control) predicted all three types of online procrastination, others (e.g., premeditation) did not influence any of these tendencies, indicating that these trait variables have separate predictive values for different types of online procrastination, and that social media procrastination, instant messaging procrastination, and shopping procrastination have distinct underlying processes and should be considered separately. Accordingly, we will first focus on boredom proneness and self-control, as all types of online procrastination were predicted by these constructs. Then, we will discuss the differences between the types of online procrastination regarding separate aspects of impulsivity. Finally, we will turn to the differences between our samples.

Boredom Proneness and Self-Control

As expected, boredom proneness was positively related to all types of online procrastination in both studies. Specifically, individuals with stronger propensity to get bored tended to use social media, instant messaging, and online shopping for procrastination more. This is in line with previous research that shows that these platforms provide an attractive alternative to procrastinate with when individuals get bored (Alblwi et al., 2019). Indeed, social media and online shopping can provide a relief from boredom (Kukar-Kinney and Close, 2010; Zolkepli and Kamarulzaman, 2015). Frequent social media use, instant messaging, and online shopping has been shown to be associated with higher boredom proneness (Matic et al., 2015; Wegmann et al., 2018). Accordingly, we found that having a higher tendency to get bored regardless of one's situation increases the likelihood to use social media and instant messaging and browse online stores rather than engage in the current task. These findings indicate that boredom proneness contributes to different types of online procrastination in addition to self-control and impulsivity.

Each type of online procrastination was also predicted by self-control. While the negative association between self-control and social media procrastination replicates prior research (Meier et al., 2016), our findings regarding procrastinatory instant messaging and shopping are novel. Specifically, individuals with lower self-control used social media, instant messaging, and online shopping for procrastination more frequently. These results support prior studies. The wish to use media is one of the most frequently experienced desires and is also amongst the desires that are most frequently surrendered to (Hofmann et al., 2012). Indeed, self-control is negatively associated with problematic media use and online shopping (Panek, 2014; Jiang et al., 2017). Having a lower capacity for "overriding prepotent responses (e.g., impulses or habits)" and refraining from acting on them results in failures of self-control (Hofmann et al., 2009, p. 165). In line with these findings, our results indicate that having lower levels of self-control capabilities increases the tendency to use social media platforms, send instant messages, and visit online stores to postpone one's tasks.

Impulsivity

Our results regarding the relationships between impulsivity and different types of online procrastination were mixed. Although the correlations between the three types of online procrastination and all impulsivity facets (except premeditation) were significant, perseverance and urgency were the only facets that predicted different types of online procrastination. Perseverance, that is, the ability to stay focused on a task until its completion, was negatively related to social media procrastination in both studies. No other facet of impulsivity was significant for procrastinatory social media use. Perseverance is associated with the capacity to hold back irrelevant thoughts as well as with trait procrastination (Bechara and Van der Linden, 2005; Rebetez et al., 2018). Individuals with lower perseverance capacities may be tempted to use their smartphones due to these irrelevant thoughts (Billieux et al., 2008b). Similarly, we found that lower perseverance increases the likelihood to use social media as tools of procrastination.

In contrast, perseverance did not influence instant messaging procrastination. While we did not have a priori expectations about other impulsivity facets than perseverance, urgency predicted procrastinatory instant messaging in Study 1, which implies different underlying mechanisms between these procrastination tendencies. Specifically, higher tendency to act impulsively when experiencing unpleasant emotional states, namely urgency, increased the likelihood to use instant messaging for procrastination. This aspect of impulsivity is associated with the number of daily SMSs sent, suggesting that, for individuals that feel like they need to pursue their impulses at once, texting can be an ideal solution when they are feeling down (Billieux et al., 2008b). Instant messaging enables communication with close others in distressing times (Cui, 2016). Indeed, students postpone assignments by first texting their friends and sharing their negative feelings (Deng, 2020). Yet, the urge to check for new online messages can contribute to daily procrastination (Meier, 2021). Our results demonstrate that higher tendency to act rash when feeling negative emotions increases procrastination with instant messaging.

Finally, we investigated procrastinatory browsing of online stores and found that perseverance and urgency were associated with it in Study 1. Specifically, lower perseverance and higher urgency simultaneously increased the tendency to visit online stores for procrastination. This is partly in line with the literature. Window-shopping can uplift consumers' moods, which is a strong motivation for online impulse purchases (Woodruffe, 1997; Sundstrom et al., 2019) and urgency is the only predictor of compulsive buying (Billieux et al., 2008a). The significance of perseverance in our results implies that, in addition to urgency, a lower capacity to remain focused on a task also increases procrastinatory shopping tendencies. Individuals who experience difficulties with staying concentrated on a task may also struggle with inhibiting task-irrelevant thoughts (e.g., a sale in a clothing store) and be more likely to browse online stores to procrastinate.

Sample Differences

Literature indicates that cultural factors affect both media use and procrastination (Mann et al., 1998; Goodrich and de Mooij, 2014). Accordingly, our two studies differed in the effects of certain predictors. Age negatively influenced all types of online procrastination for the Germans, such that, younger Germans were more likely to use the Internet for procrastination. This is in line with Beutel et al. (2016), who found that trait procrastination decreased with age across German samples. Although these behaviors did decrease with age with our Turkish participants, it did not predict procrastination, which replicates past findings in Turkey (Bekleyen, 2017). Alternatively, the restricted age range and variance in the Turkish sample might have prevented age from becoming a significant predictor, although in both samples younger individuals used social media, instant messaging, and online shopping for procrastination more than their older counterparts.

We further found that, for the Germans but not the Turkish, perseverance predicted shopping procrastination and urgency predicted instant messaging and shopping procrastination. These cultural differences in the effect of impulsivity resemble prior research, which indicate that culture has an influence on actual and problematic (online) shopping behaviors and the frequency of visiting online stores (Gong, 2009; Baron and af Segerstad, 2010). Overall, for the Turkish, impulsivity does not make any significant contribution while predicting procrastinatory instant messaging and shopping as it does for the Germans.

Limitations

One limitation of this study is its correlational design, which precludes definite causal inferences. Longitudinal studies should clarify the directionality between boredom proneness and media use. Moreover, we did not differentiate between devices (e.g., smartphone, computer) that were used when accessing these platforms, although social media sites are accessed increasingly more through smartphones and tablets compared to personal computers (Droesch, 2019). Future studies on online procrastination could delve into possible differences between the mobile devices and computers as tools of procrastination. Furthermore, as some studies indicate that social media can be used both to escape unpleasant life situations as well as to procrastinate (Meier et al., 2018), the role of escapism in online procrastination should be further explored. Finally, we need to consider the impact of the COVID-19 pandemic, during which most people worked from home. Our results might also be influenced by factors such as lower structure and motivation (Melgaard et al., 2022).

CONCLUSION

Online procrastination is an increasingly common phenomenon. Literature has investigated the uses of general media, Facebook, and smartphones for procrastination. The purpose of this paper was to better understand the connections between several personality traits and types of online procrastination. Accordingly, we examined the influence of boredom proneness, self-control, and impulsivity on procrastinatory social media use, instant messaging, and online shopping tendencies. Our results show that, in addition to self-control, boredom proneness is especially predictive of online procrastination.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository and accession number can be found at: https://osf.io/czbka/.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Ethics Committee of the Department of Computer Science and Applied Cognitive Science, University of Duisburg-Essen, Duisburg, Germany. The participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

CS and OB designed the study. CS organized data collection, performed the statistical analysis, and wrote the first draft of the manuscript. CS and OB revised the manuscript and approved the submitted version. Both authors contributed to the article and approved the submitted version.

FUNDING

CS was a scholarship recipient under the Graduate School Scholarship Programme from the Deutscher Akademischer Austauschdienst (DAAD) funding ID 57516591.

REFERENCES

Achtziger, A., Hubert, M., Kenning, P., Raab, G., and Reisch, L. (2015). Debt out of control: the links between self-control, compulsive buying, and real debts. *J. Econ. Psychol.* 49, 141–149. doi: 10.1016/j.joep.2015.04.003

Alblwi, A., Stefanidis, A., Phalp, K., and Ali, R. (2019). "Procrastination on social networks: types and triggers," in *Proceedings of the 6th International Conference on Behavioral, Economic and Socio-Cultural Computing*, Beijing.

Balkis, M., and Duru, E. (2019). Procrastination and rational/irrational beliefs: a moderated mediation model. *J. Ration. Emot. Cogn. Behav. Ther.* 37, 299–315. doi: 10.1007/s10942-019-00314-6

Baron, N. S., and af Segerstad, Y. H. (2010). Cross-cultural patterns in mobile-phone use: public space and reachability in Sweden, the USA and Japan. *New Media Soc.* 12, 13–34. doi: 10.1177/1461444809355111

Baumeister, R. F., Bratslavsky, E., Muraven, M., and Tice, D. M. (1998). Ego depletion: is the active self a limited resource? *J. Pers. Soc. Psychol.* 74, 1252–1265. doi: 10.1037/0022-3514.74.5.1252

Bayer, J. B., Dal Cin, S., Campbell, S. W., and Panek, E. (2016). Consciousness and self-regulation in mobile communication. *Hum. Commun. Res.* 42, 71–97. doi: 10.1111/hcre.12067

Bechara, A., and Van der Linden, M. (2005). Decision-making and impulse control after frontal lobe injuries. *Curr. Opin. Neurol.* 18, 734–739. doi: 10.1097/01.wco.0000194141.56429.3c

Bekleyen, N. (2017). Understanding the academic procrastination attitude of language learners in Turkish universities. *Educ. Res. Rev.* 12, 108–115.

Bench, S. W., and Lench, H. C. (2013). On the function of boredom. *Behav. Sci.* 3, 459–472. doi: 10.3390/bs3030459

Bertrams, A., and Dickhäuser, O. (2009). Messung dispositioneller selbstkontrollkapazität: eine deutsche adaptation der kurzform der self-control scale. *Diagnostica* 55, 2–10. doi: 10.1026/0012-1924.55.1.2

Beswick, G., Rothblum, E. D., and Mann, L. (1988). Psychological antecedents of student procrastination. *Aust. Psychol.* 23, 207–217. doi: 10.1080/00050068808255605

Beutel, M. E., Klein, E. M., Aufenanger, S., Brahler, E., Dreier, M., Muller, K. W., et al. (2016). Procrastination, distress and life satisfaction across the age range-a german representative community study. *PLoS One* 11:e0148054. doi: 10.1371/journal.pone.0148054

Billieux, J., Van Der Linden, M., and Rochat, L. (2008b). The role of impulsivity in actual and problematic use of the mobile phone. *Appl. Cogn. Psychol.* 22, 1195–1210. doi: 10.1002/acp.1429

Billieux, J., Rochat, L., Rebetez, M. M. L., and Van der Linden, M. (2008a). Are all facets of impulsivity related to self-reported compulsive buying behavior? *Pers. Individ. Diff.* 44, 1432–1442. doi: 10.1016/j.paid.2007.12.011

Biolcati, R., Mancini, G., and Trombini, E. (2018). Proneness to boredom and risk behaviors during adolescents' free time. *Psychol. Rep.* 121, 303–323. doi: 10.1177/0033294117724447

Blight, M. G., Ruppel, E. K., and Schoenbauer, K. V. (2017). Sense of community on twitter and instagram:exploring the roles of motives and parasocial relationships. *Cyberpsychol. Behav. Soc. Netw.* 20, 314–319. doi: 10.1089/cyber.2016.0505

Blunt, A. K., and Pychyl, T. A. (1998). Volitional action and inaction in the lives of undergraduate students: state orientation, procrastination and proneness to boredom. *Pers. Individ. Diff.* 24, 837–846. doi: 10.1016/S0191-8869(98)00018-X

Bozaci, I. (2020). The effect of boredom proneness on smartphone addiction and impulse purchasing: a field study with young consumers in turkey. *J. Asian Finance Econ. Bus.* 7, 509–517. doi: 10.13106/jafeb.2020.vol7.no7.509

Cui, D. (2016). Beyond "connected presence": multimedia mobile instant messaging in close relationship management. *Mobile Media Commun.* 4, 19–36. doi: 10.1177/2050157915583925

Deng, L. P. (2020). Laptops and mobile phones at self-study time: examining the mechanism behind interruption and multitasking. *Aust. J. Educ. Technol.* 36, 55–67. doi: 10.14742/ajet.5048

Droesch, B. (2019). *More Than Half of US Social Network Users Will Be Mobile-Only in 2019*. Available online at: https://www.emarketer.com/content/more-than-half-of-social-network-users-will-be-mobile-only-in-2019 [Accessed June 7, 2021].

Eastwood, J. D., Frischen, A., Fenske, M. J., and Smilek, D. (2012). The unengaged mind: defining boredom in terms of attention. *Perspect. Psychol. Sci.* 7, 482–495. doi: 10.1177/1745691612456044

Eurostat (2022). *E-Commerce Statistics For Individuals*. Luxembourg City: Eurostat.

Farmer, R., and Sundberg, N. D. (1986). Boredom proneness-the development and correlates of a new scale. *J. Pers. Assess.* 50, 4–17. doi: 10.1207/s15327752jpa5001_2

Ferrari, J. R. (2000). Procrastination and attention:factor analysis of attention deficit, boredomness, intelligence, self-esteem, and task delay frequencies. *J. Soc. Behav. Pers.* 15, 185–196.

Fujita, K. (2011). On conceptualizing self-control as more than the effortful inhibition of impulses. *Pers. Soc. Psychol. Rev.* 15, 352–366. doi: 10.1177/1088868311411165

Gong, W. (2009). National culture and global diffusion of business-to-consumer e-commerce. *Cross Cult. Manage.* 16, 83–101. doi: 10.1108/13527600910930059

Goodrich, K., and de Mooij, M. (2014). How 'social' are social media? A cross-cultural comparison of online and offline purchase decision influences. *J. Mark. Commun.* 20, 103–116. doi: 10.1080/13527266.2013.797773

Güner, H., Okan, N., and Kardaş, S. (2021). Kısa can sikintisi eğilimi ölçeğinin türkçe'ye uyarlanmasi ve psikometrik yönden incelenmesi. *Marmara Üniv. Atatürk Eğitim Fakültesi Eğitim Bilimleri Dergisi* 53, 326–341. doi: 10.15285/maruaebd.797235

Hofmann, W., Friese, M., and Strack, F. (2009). Impulse and self-control from a dual-systems perspective. *Perspect. Psychol. Sci.* 4, 162–176. doi: 10.1111/j.1745-6924.2009.01116.x

Hofmann, W., Reinecke, L., and Meier, A. (2017). "Of sweet temptations and bitter aftertaste: self-control as a moderator of the effects of media use on well-being," in *The Routledge Handbook Of Media Use And Well-Being: International Perspectives On Theory And Research On Positive Media Effects*, eds L. Reinecke and M. B. Oliver (Abingdon: Routledge), 211–222.

Hofmann, W., Vohs, K. D., and Baumeister, R. F. (2012). What people desire, feel conflicted about, and try to resist in everyday life. *Psychol. Sci.* 23, 582–588. doi: 10.1177/0956797612437426

Jiang, Z., Zhao, X., and Li, C. (2017). Self-control predicts attentional bias assessed by online shopping-related Stroop in high online shopping addiction tendency college students. *Compr. Psychiatry* 75, 14–21. doi: 10.1016/j.comppsych.2017.02.007

Keye, D., Wilhelm, O., and Oberauer, K. (2009). Structure and correlates of the german version of the brief UPPS impulsive behavior scales. *Eur. J. Psychol. Assess.* 25, 175–185. doi: 10.1027/1015-5759.25.3.175

Klingsieck, K. B. (2013). Procrastination when good things don't come to those who wait. *Eur. Psychol.* 18, 24–34. doi: 10.1027/1016-9040/a000138

Koessmeier, C., and Büttner, O. B. (2021). Why are we distracted by social media? Distraction situations and strategies, reasons for distraction, and individual differences. *Front. Psychol.* 12:711416. doi: 10.3389/fpsyg.2021.711416

Kukar-Kinney, M., and Close, A. G. (2010). The determinants of consumers' online shopping cart abandonment. *J. Acad. Mark. Sci.* 38, 240–250. doi: 10.1007/s11747-009-0141-5

Lavoie, J. A. A., and Pychyl, T. A. (2001). Cyberslacking and the procrastination superhighway: a web-based survey of online procrastination, attitudes, and emotion. *Soc. Sci. Comput. Rev.* 19, 431–444. doi: 10.1177/089443930101900403

Levine, L. E., Waite, B. M., and Bowman, L. L. (2013). Use of instant messaging predicts self-report but not performance measures of inattention, impulsiveness, and distractibility. *Cyberpsychol. Behav. Soc. Netw.* 16, 898–903. doi: 10.1089/cyber.2012.0504

Li, C. K. W., Holt, T. J., Bossler, A. M., and May, D. C. (2016). Examining the mediating effects of social learning on the low self-control-cyberbullying relationship in a youth sample. *Deviant Behav.* 37, 126–138. doi: 10.1080/01639625.2014.1004023

Mann, L., Radford, M., Burnett, P., Ford, S., Bond, M., Leung, K., et al. (1998). Cross-cultural differences in self-reported decision-making style and confidence. *Int. J. Psychol.* 33, 325–335. doi: 10.1080/002075998400213

Martarelli, C. S., Bertrams, A., and Wolff, W. (2020). A personality trait-based network of boredom, spontaneous and deliberate mind-wandering. *Assessment* 28, 1915–1931. doi: 10.1177/1073191120936336

Martin, M., Sadlo, G., and Stew, G. (2006). The phenomenon of boredom. *Qual. Res. Psychol.* 3, 193–211. doi: 10.1191/1478088706qrp066oa

Martínez-López, F. J., Pla-García, C., Gázquez-Abad, J. C., and Rodríguez-Ardura, I. (2016). Hedonic motivations in online consumption behaviour. *Int. J. Bus. Environ.* 8, 121–151. doi: 10.1504/IJBE.2016.076628

Matic, A., Pielot, M., and Oliver, N. (2015). "Boredom-computer interaction: boredom proneness and the use of smartphone," in *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, Osaka, doi: 10.1145/2750858.2807530

Meier, A. (2021). Studying problems, not problematic usage: do mobile checking habits increase procrastination and decrease well-being? *Mobile Media Commun.* 10, 272–279. doi: 10.1177/20501579211029326

Meier, A., Meltzer, C. E., and Reinecke, L. (2018). "Coping with stress or losing control? Facebook-induced strains among emerging adults as a consequence of escapism versus procrastination," in *Youth and Media: Current Perspectives on Media Use and Effects*, eds R. Kühne, S. E. Baumgartner, T. Koch, and M. Hofer (Baden-Baden: Nomos Verlagsgesellschaft mbH and Co. KG), 167–186. doi: 10.5771/9783845280455-167

Meier, A., Reinecke, L., and Meltzer, C. E. (2016). "Facebocrastination"? Predictors of using Facebook for procrastination and its effects on students' well-being. *Comput. Hum. Behav.* 64, 65–76. doi: 10.1016/j.chb.2016.06.011

Melgaard, J., Monir, R., Lasrado, L. A., and Fagerstrøm, A. (2022). Academic procrastination and online learning during the COVID-19 pandemic. *Proc. Comput. Sci.* 196, 117–124. doi: 10.1016/j.procs.2021.11.080

Mercer-Lynn, K. B., Bar, R. J., and Eastwood, J. D. (2014). Causes of boredom: the person, the situation, or both? *Pers. Individ. Diff.* 56, 122–126. doi: 10.1016/j.paid.2013.08.034

Montag, C., Błaszkiewicz, K., Sariyska, R., Lachmann, B., Andone, I., Trendafilov, B., et al. (2015). Smartphone usage in the 21st century: who is active on WhatsApp? *BMC Res. Notes* 8:331. doi: 10.1186/s13104-015-1280-z

Nebioglu, M., Konuk, N., Akbaba, S., and Eroglu, Y. (2012). The investigation of validity and reliability of the turkish version of the brief self-control scale. *Bull. Clin. Psychopharmacol.* 22, 340–351. doi: 10.5455/bcp.20120911042732

Orchard, L. J., Fullwood, C., Galbraith, N., and Morris, N. (2014). Individual differences as predictors of social networking. *J. Comput. Med. Commun.* 19, 388–402. doi: 10.1111/jcc4.12068

Panek, E. (2014). Left to their own devices college students' "Guilty pleasure" media use and time management. *Commun. Res.* 41, 561–577. doi: 10.1177/0093650213499657

Parsons, A. G. (2002). Non-functional motives for online shoppers:why we click. *J. Consumer Mark.* 19, 380–392. doi: 10.1108/07363760210437614

Pychyl, T. A., and Sirois, F. M. (2016). "Procrastination, emotion regulation, and well-being," in *Procrastination, Health, And Well-Being*, eds F. M. Sirois and T. A. Pychyl (Amsterdam: Elsevier Science), 163–188.

Quan-Haase, A., and Young, A. L. (2010). Uses and gratifications of social media: a comparison of facebook and instant messaging. *Bull. Sci. Technol. Soc.* 30, 350–361. doi: 10.1177/0270467610380009

Rebetez, M. M. L., Rochat, L., Barsics, C., and Van der Linden, M. (2018). Procrastination as a self-regulation failure:the role of impulsivity and intrusive thoughts. *Psychol. Rep.* 121, 26–41. doi: 10.1177/0033294117720695

Reinecke, L., and Hofmann, W. (2016). Slacking off or winding down? An experience sampling study on the drivers and consequences of media use for recovery versus procrastination. *Hum. Commun. Res.* 42, 441–461. doi: 10.1111/hcre.12082

Reinecke, L., Hartmann, T., and Eden, A. (2014). The guilty couch potato: the role of ego depletion in reducing recovery through media use. *J. Commun.* 64, 569–589. doi: 10.1111/jcom.12107

Schnauber-Stockmann, A., Meier, A., and Reinecke, L. (2018). Procrastination out of habit? The role of impulsive versus reflective media selection in procrastinatory media use. *Media Psychol.* 21, 640–668. doi: 10.1080/15213269.2018.1476156

Sindermann, C., Elhai, J. D., and Montag, C. (2020). Predicting tendencies towards the disordered use of Facebook's social media platforms: on the role of personality, impulsivity, and social anxiety. *Psychiatry Res.* 285:112793. doi: 10.1016/j.psychres.2020.112793

Singh, M. (2020). *WhatsApp Is Now Delivering Roughly 100 Billion Messages A Day*. Available online at: https://techcrunch.com/2020/10/29/whatsapp-is-now-delivering-roughly-100-billion-messages-a-day/ [Accessed May 17, 2021].

Sirois, F., and Pychyl, T. (2013). Procrastination and the priority of short-term mood regulation: consequences for future self. *Soc. Pers. Psychol. Compass* 7, 115–127. doi: 10.1111/spc3.12011

Steel, P. (2007). The nature of procrastination: a meta-analytic and theoretical review of quintessential self-regulatory failure. *Psychol. Bull.* 133, 65–94. doi: 10.1037/0033-2909.133.1.65

Steel, P. (2010). Arousal, avoidant and decisional procrastinators: do they exist? *Pers. Individ. Diff.* 48, 926–934. doi: 10.1016/j.paid.2010.02.025

Struk, A. A., Carriere, J. S. A., Cheyne, J. A., and Danckert, J. (2017). A short boredom proneness scale: development and psychometric properties. *Assessment* 24, 346–359. doi: 10.1177/1073191115609996

Sundstrom, M., Hjelm-Lidholm, S., and Radon, A. (2019). Clicking the boredom away-exploring impulse fashion buying behavior online. *J. Retailing Consumer Serv.* 47, 150–156. doi: 10.1016/j.jretconser.2018.11.006

Svartdal, F., Pfuhl, G., Nordby, K., Foschi, G., Klingsieck, K. B., Rozental, A., et al. (2016). On the measurement of procrastination: comparing two scales in six european countries. *Front. Psychol.* 7:1307. doi: 10.3389/fpsyg.2016.01307

Tam, K. Y. Y., van Tilburg, W. A. P., and Chan, C. S. (2021). What is boredom proneness? A comparison of three characterizations. *J. Pers.* 89, 831–846. doi: 10.1111/jopy.12618

Tangney, J. P., Baumeister, R. F., and Boone, A. L. (2004). High self-control predicts good adjustment, less pathology, better grades, and interpersonal success. *J. Pers.* 72, 271–324. doi: 10.1111/j.0022-3506.2004.00263.x

Troll, E. S., Friese, M., and Loschelder, D. D. (2021). How students' self-control and smartphone-use explain their academic performance. *Comput. Hum. Behav.* 117:106624. doi: 10.1016/j.chb.2020.106624

Vodanovich, S. J., and Rupp, D. E. (1999). Are procrastinators prone to boredom? *Soc. Behav. Pers.* 27, 11–16. doi: 10.2224/sbp.1999.27.1.11

Wegmann, E., Ostendorf, S., and Brand, M. (2018). Is it beneficial to use Internet-communication for escaping from boredom? Boredom proneness interacts with cue-induced craving and avoidance expectancies in explaining symptoms of Internet-communication disorder. *PLoS One* 13:e0195742. doi: 10.1371/journal.pone.0195742

Westgate, E. C., and Wilson, T. D. (2018). Boring thoughts and bored minds: the MAC model of boredom and cognitive engagement. *Psychol. Rev.* 125, 689–713. doi: 10.1037/rev0000097

Whiteside, S. P., and Lynam, D. R. (2001). The five factor model and impulsivity: using a structural model of personality to understand impulsivity. *Pers. Individ. Diff.* 30, 669–689. doi: 10.1016/S0191-8869(00)00064-7

Wijaya, H. E., and Tori, A. R. (2018). Exploring the role of self-control on student procrastination. *Int. J. Res. Couns. Educ.* 1:13. doi: 10.24036/003za0002

Woodruffe, H. R. (1997). Compensatory consumption: Why women go shopping when they're fed up and other stories. *Mark. Intell. Plan.* 15, 325–334. doi: 10.1108/02634509710193172

Wypych, M., Matuszewski, J., and Dragan, W. L. (2018). Roles of impulsivity, motivation, and emotion regulation in procrastination-path analysis and comparison between students and non-students. *Front. Psychol.* 9:891. doi: 10.3389/fpsyg.2018.00891

Yargıç, I., Ersoy, E., and Oflaz, S. B. (2011). Measuring impulsivity of psychiatric patients using UPPS impulsive behavior scale. *Bull. Clin. Psychopharmacol.* 21, 139–146. doi: 10.5455/bcp.20110706024203

Zolkepli, I. A., and Kamarulzaman, Y. (2015). Social media adoption: the role of media needs and innovation characteristics. *Comput. Hum. Behav.* 43, 189–209. doi: 10.1016/j.chb.2014.10.050

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

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 18 August 2022
DOI 10.3389/fpsyg.2022.953232

OPEN ACCESS

EDITED BY
Martina Ziefle,
RWTH Aachen University, Germany

REVIEWED BY
André Calero Valdez,
University of Lübeck, Germany
Umair Akram,
RMIT University, Vietnam
Vera Gelashvili,
Rey Juan Carlos University, Spain
Aditya Halim Perdana Kusuma Putra,
Universitas Muslim Indonesia, Indonesia
Mingyue Fan,
Jiangsu University, China

*CORRESPONDENCE
Luyao Liu
liuluyao2019@163.com

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

RECEIVED 26 May 2022
ACCEPTED 01 August 2022
PUBLISHED 18 August 2022

CITATION
Xie X and Liu L (2022) Exploring the antecedents of trust in electronic word-of-mouth platform: The perspective on gratification and positive emotion.
Front. Psychol. 13:953232.
doi: 10.3389/fpsyg.2022.953232

Exploring the antecedents of trust in electronic word-of-mouth platform: The perspective on gratification and positive emotion

Xuemei Xie and Luyao Liu*

School of Economics and Management, Beijing University of Posts and Telecommunications, Beijing, China

Frequent human-media interaction *via* the electronic word-of-mouth (e-wom) platform, trust is acknowledged as an ongoing challenge. This study aimed to understand users' trust in the e-wom platform based on uses and gratifications theory and stimulus-organism-response (S-O-R) paradigm. Utilitarian gratification (perceived information quality and perceived privacy protection) was regarded as stimulus, social gratification (sense of social belonging and sense of self-esteem) and positive emotion as organism, and platform trust as response. Data was acquired from 268 users in China using a questionnaire survey, and the PLS-SEM was used to further analyze the results. The results indicated that there is a hierarchy relationship between types of gratifications. That is, utilitarian gratification is a premise of social gratification. Moreover, sense of self-esteem and positive emotion have a mediating effect between perceived information quality and platform trust. Sense of social belonging and positive emotion have a mediating effect between perceived privacy protection and platform trust. This study not only broadened trust between human and media, but also purposed a hierarchy relationship of different types of gratifications in e-wom platform.

KEYWORDS

e-wom platform, platform trust, gratification, positive emotion, S-O-R paradigm

1. Introduction

As new platform technologies are introduced, the topic of how to improve trust takes on a new dimension. The e-wom platform provides a new set of options for users to share and get information about products and services (Hu et al., 2019; Lin et al., 2020). Since its introduction in 2013, XIAOHONGSHU has grown to become China's most popular e-wom platform. On the XIAOHONGSHU platform, over 300 million people share their purchasing and life experiences, allowing other users to make better decisions. According to a recent report, the number of content creators on the XIAOHONGSHU platform reached 43 million by March 2021, and the number of notes surpassed 300 million.

The number of monthly active users hit 282 million in October 2021. The above data reveal frequently increased human-media interaction on the e-wom platform. In the digital world, trust is the base between the platform and users, as well as the cornerstone of the platform's survival (Li and Lin, 2021). Therefore, it is meaningful to explore various antecedents that influence trust in the e-wom platform.

Prior studies mainly explored users' behaviors on e-wom information, such as adoption, engage, and spread. Some researchers explored why people adopt e-wom on social networking sites based on the attachment theory (Park et al., 2019). Attachment avoidance, attachment anxiety, and interaction effects are the antecedents of e-wom adoption. Some researchers investigated how former customers engage in e-wom (Azer and Ranaweera, 2022). Some researchers investigated various motivational factors that influence SNS users' e-wom intention (Chai et al., 2022). They found intrinsic motivational factors embracing altruism, self-efficacy, and self-expression universally influence SNS users' e-wom intention. However, empirical study on the antecedents of trust in the e-wom platform is still lacking.

Trust is not a rational cognitive process (Yuan et al., 2018). Emotional states play a significant role. In this study, we explored key factors influencing users' trust in the e-wom platform, which include various gratifications and positive emotion. Especially, based on users and gratifications (UG) theory, this study proposed utilitarian gratification (perceived information quality and perceived privacy protection) and social gratification (sense of social belonging and sense of self-esteem). Combined with stimulus-organism-response (S-O-R) paradigm, utilitarian gratification was considered as stimulus, social gratification, and positive emotion as organism, and platform trust as response. This study found that utilitarian gratification is a premise of social gratification. To some extent, social gratification and positive emotion have a mediating effect between utilitarian gratification and platform trust. This study extended UG theory and S-O-R paradigm in the context of the e-wom platform. We broadened e-wom research, especially regarding how users trust in the e-wom platform. a hierarchy relationship exists in different types of gratifications.

2. Theoretical foundation and hypotheses

2.1. Different types of gratifications

Uses and gratifications theory, was first created in the research field of mass media like radio, newspaper, television. Its goal is to figure out what motivates people to use certain types of media (Leung and Wei, 2000). Meanwhile, it examines why people choose one form of media over another in order to gratify a variety of needs (Katz et al., 1974). With rapid information system, today it is widely used in the research field of social media to better understand psychological state of users.

Present studies have classified gratifications obtained when using various social media. Some researchers investigated the determinants of continuance intention toward SNSs (Chang, 2018). They found perceived gratifications including information gratification, emotional gratification, and social gratification. Some researchers examined the effects of different gratifications on the continuance intention of using WeChat in China (Gan and Li, 2018). They identified four types of gratifications, namely hedonic gratification, social gratification, utilitarian gratification, and technology gratification. Some researchers determined the impact of gratifications and emotional state on users' adoption and continuance intention in Weibo (Gogan et al., 2016). They explored users' gratifications, namely hedonic gratification (entertaining value), social gratification (social value and social participation), and utilitarian gratification (information consumption, utilitarian value, and content participation). Some researchers examined continuance intention with live-streaming services based on UG theory (Hsu and Lin, 2021). They determined three gratifications, namely entertainment gratification, informativeness gratification, and sociability gratification. Some researchers examined the antecedents of grocery purchase behavior, and identified three types of gratification, including utilitarian gratification, hedonic gratification, and experiential gratification (Kim, 2021).

Present studies pay more attention on the parallel relationship among different gratifications, instead of seeking other relationship. To fill this research gap, this study intended to employ perspectives of hierarchy of needs to identify hierarchy relationship of different types of gratifications. Based on UG theory and characteristics of the e-wom platform, we purposed utilitarian gratification and social gratification. Utilitarian gratification includes perceived information quality and perceived privacy protection. Social gratification includes sense of social belonging and sense of self-esteem.

The definition of perceived information quality refers to the correctness and completeness of website information as it relates to products and transactions (Kim et al., 2008). Users' perceptions of information quality have a beneficial impact on their willingness to participate actively in platform communication and engagement (Lu et al., 2011). As a result, we predicted that if the e-wom platform continues to provide meaningful and useful information, users' sense of social belonging will improve dramatically. Therefore, we put forward the following hypothesis:

H1: Perceived information quality has a positive impact on sense of social belonging.

Self-concept is a comprehensive view formed by individuals' cognition of themselves in various aspects. Self-esteem is typically influenced by environmental cues, information (evaluation and expectation) from influential figures in

societal structure, and sense of competence and efficiency that individuals have experienced (Brockner, 1988). When people create a sense of importance and worth for themselves, they develop self-esteem (Pan et al., 2012). Specific to the e-wom platform, on the one hand, when users contribute information about products or services, and others accept and recognize their knowledge, they believe that they are capable and valuable. Users, on the other hand, will feel appreciated and so boost their self-esteem if they collect knowledge on the platform and others actively assist them. According to a research paper from organization behavior (Zheng et al., 2017), knowledge sharing has a positive effect on organization-based self-esteem. Thus, when users on the e-wom platform provide or acquire high quality information, namely, they obtain information gratification, their sense of self-esteem would improve dramatically. Therefore, we put forward the following hypothesis:

H2: Perceived information quality has a positive impact on sense of self-esteem.

Individuals' activity in social networks with high and transparent sociability can be witnessed by numerous others, making it impossible to hide (Livingstone, 2008). Privacy issues may arise if too much information is shared and received by too many persons (Schwartz, 1968). Previous research has found that overly apparent privacy policies can deter users from sharing content (Brandtzæg et al., 2010). Privacy protection is an important aspect in the growth of social media, as it encourages social contact between users (Sapuppo and Seet, 2015). We propose that, because of the platform's openness, users not only communicate utilitarian knowledge but also emotional experiences. As a result, privacy protection is at the heart of social interaction. This study suggested that privacy protection creates a secure interactive environment, in which users strive to form social relationships with one another and increase their sense of social belonging after obtaining technology gratification. Thus, we put forward the following hypothesis:

H3: Perceived privacy protection has a positive impact on sense of social belonging.

Today, the number of weak-tie contacts has explored *via* online social media. Individual performance is divided into on-stage and off-stage, according to Goffman (2016). In weak ties, individuals are more likely to execute impression management, which means they are more likely to project a positive image in front of strangers (Luo and Cong, 2015). On the e-wom platform which characterized by weak ties, self-image management corresponds to "off-stage performance," which is not visible to on-stage acquaintances. If the e-wom platform has adequate privacy protection, users prefer to perform self-image management on this platform, boosting individuals' sense of self-esteem. We predicted that the cornerstone of social gratification is privacy protection. Therefore, we purposed the following hypothesis:

H4: Perceived privacy protection has a positive impact on sense of self-esteem.

2.2. Social gratification and positive emotion

Positive emotion is linked to the satisfaction of a certain need, which is frequently accompanied by a good subjective experience, and can boost an individual's excitement and ability to participate in activities (Meng, 1989).

The augmentation of emotional value offered by self-expression and relationships with others through the platform, so that users are more emotionally linked to the platform, is referred to as a sense of social belonging (Hsu and Lin, 2016). Previous research has shown that using social networking sites can successfully lessen loneliness and increase pleasant emotions (Huang et al., 2016). Users on the e-wom platform have a sense of social belonging through social interaction such as commenting and messaging. Sense of social belonging can help people feel better. Therefore, we put forward the following hypothesis:

H5: Sense of social belonging has a positive impact on positive emotion.

Individual self-worth and importance are reflected in self-esteem, while positive emotion represents an individual's emotional condition, such as happiness. There is a relationship between cognition and emotion (Lavy and Littman-Ovadia, 2017). Thus, we suggest that individual self-esteem as a cognition can directly influence positive emotion. According to attribution theory (Heider, 1958), individuals' self-esteem can be attributed internally and externally. Users hint at their own taste and social standing by flaunting their wonderful lives, which is known as internal attribution. External attribution is that the e-wom platform has a mutual respect culture. Because the majority of users are young on the e-wom platform, have a high level of education and quality. Previous study in the field of psychology suggests that self-esteem has a positive influence on position emotion (Liu et al., 2018). Thus, we predicted that on the e-wom platform, sense of self-esteem can enhance positive emotional state. We put forward the following hypothesis:

H6: Sense of self-esteem has a positive impact on positive emotion.

2.3. Positive emotion and trust

Positive emotion refers to the emotion with pleasant feeling is generated when individuals are stimulated by internal and external environment, and meet their own needs (Fredrickson et al., 2008). According to the broaden-and-build theory, positive emotion extends the scope of people's attention and thought-action repertoires, and create enduring

personal resources including social connections and intellectual resources, as well as flourish people's mental health (Fredrickson and Losada, 2005). In exchange for favorable expectations of another's intentions or actions, trust is a psychological state characterized by the willingness to tolerate vulnerability (Rousseau et al., 1998). Meanwhile, trust is defined by dynamic and situational social psychological phenomena, and it has a complicated link with people's psychology, with good feeling being one of the most essential components (Wang and Lian, 1998). The relationship between positive emotions and trust has been discovered through interpersonal trust studies. According to affect-as-information theory, individuals use their affective states as information when making decisions (Schwarz and Clore, 1983). Happiness and gratitude, both positive emotions, improve interpersonal trust, but rage, a negative emotion, diminishes interpersonal trust (Dunn and Schweitzer, 2005).

Further, this study extended trust from level of interpersonal relationship to human and media. When people make evaluative judgments about trust, specific emotions influence subsequent judgments (Dunn and Schweitzer, 2005). In other words, different emotions can provide information sources. Trust is not automatic, but based on positive emotions. Positive emotion can boost people's positive perceptions of their risk partner, leading to greater positive decision-making, including trust (Bless and Fiedler, 2006). We predicted that when users place their trust in the e-wom platform, positive emotion offers evidence about trust judgment. Therefore, we put forward the following hypothesis:

H7: Positive emotion has a positive impact on platform trust.

2.4. The stimulus-organism-response paradigm

Given that utilitarian gratification is a crucial stimulus for the platform trust, users' social gratification and positive emotion dominate their final trust decision (response). The justification for using the stimulus-organism-response (S-O-R) paradigm as the theoretical lens to examine how types of gratifications and positive emotion influence trust in the e-wom platform is as follow. First, prior researchers (Yuan et al., 2020, 2021) applied the S-O-R paradigm to predict users' attitude, such as loyalty in social media. Second, its theoretical justification of examining utilitarian gratifications as stimuli and its capability of evaluating the role that users' emotional perceptions (social gratification and positive emotion) to utilitarian gratification plays in users' trust in the e-wom platform.

S-O-R paradigm to specify mediating processes in an organism that transmit a stimulus to a response. The term organism refers to the internal processes and structures intervening between stimuli and the final responses emitted. The intervening processes and structures consist of perceptual, physiological, feeling, and thinking activities. Response pertains to psychological reactions such as attitudinal and behavioral reactions. According to the S-O-R paradigm, users' utilitarian gratification (stimulus) may affect users' emotional perceptions (social gratification and positive emotion) (organism), which in turn may influence users' trust in the platform (response).

Therefore, we proposed the following hypotheses:

H8a: Sense of social belonging and positive emotion play chain double-mediation effects on the relationship between perceived information quality and platform trust.

H8b: Sense of self-esteem and positive emotion play chain double-mediation effects on the relationship between perceived information quality and platform trust.

H9a: Sense of social belonging and positive emotion play chain double-mediation effects on the relationship between perceived privacy protection and platform trust.

H9b: Sense of self-esteem and positive emotion play chain double-mediation effects on the relationship between perceived privacy protection and platform trust.

Theoretical model is as shown in Figure 1.

3. Research methodology

3.1. Measurement

Multiple items are used to measure all constructs, which are gathered in the survey using a five-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). The constructs are mostly adapted from earlier studies, but have been modified to request data on the e-wom platform. To ensure content validity, we invite experts to modify these items and users of the e-wom platform to do a pre-test.

The measurement of perceived information quality is mainly referenced to Kim et al. (2008). The measurement of perceived privacy protection is mainly referenced to Kim et al. (2008). The measurement of sense of social belonging is mainly referenced to Lin (2008). The measurement of sense of self-esteem is mainly referenced to Rosenberg (1965). The measurement of positive emotion is mainly referenced to Fredrickson (2013). Finally, the measurement of platform trust is mainly referenced to Suh and Han (2003). Items and sources are as shown in Table 1.

3.2. Data collection and sample description

We disseminate the online questionnaire using Wenjuanxing, a professional online questionnaire platform. Only respondents who had used the e-wom platform before the poll were eligible to participate, in order to

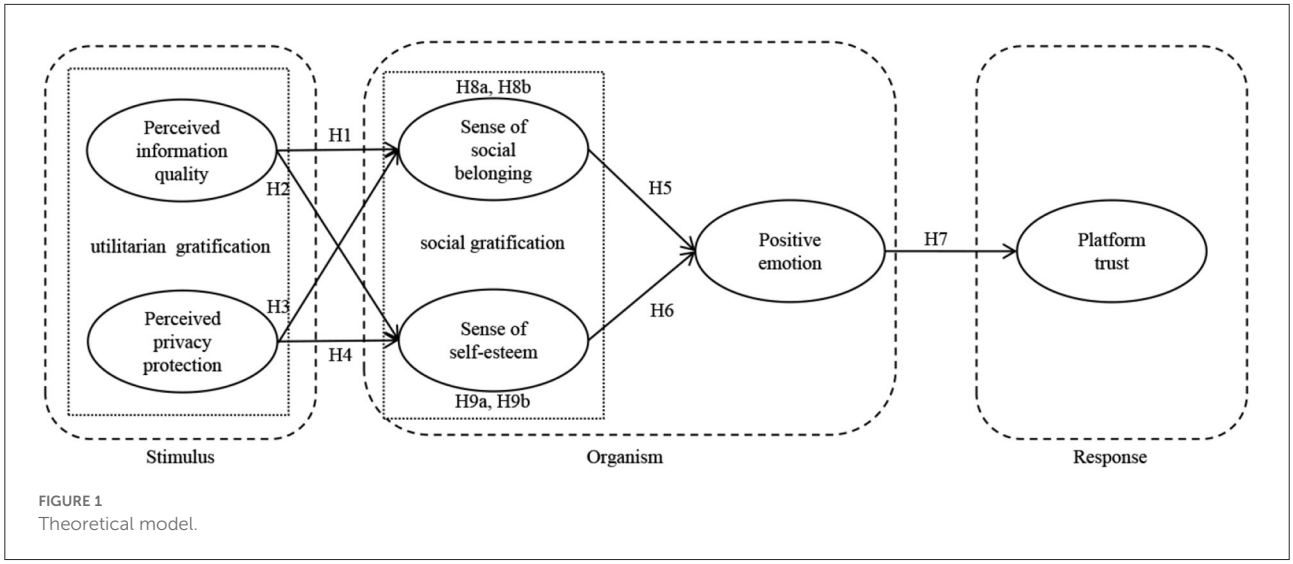


FIGURE 1
Theoretical model.

TABLE 1 Item and source.

Construct		Item	References
Perceived information quality (PIQ)	PIQ1	I think the word-of-mouth information within this platform is reliable	Kim et al., 2008
	PIQ2	I think the word-of-mouth information within this platform is useful.	
	PIQ3	I think the word-of-mouth information within this platform is of high quality.	
Perceived privacy protection (PPP)	PPP1	I think this platform will not use my personal information for other purposes without my authorization.	Kim et al., 2008
	PPP2	I think this platform will not share my personal information with other entities without my authorization.	
	PPP3	I think unauthorized persons have not access to my personal information.	
Sense of social belonging (SSB)	SB1	I enjoy being a member of this platform.	Lin, 2008
	SB2	I am very committed to this platform.	
	SB3	I enjoy this platform that has a high level of morale.	
	SB4	I feel a strong sense of social belonging to this platform.	
Sense of self-esteem (SSE)	SE1	I more recognize myself in the process of using the e-wom platform.	Rosenberg, 1965
	SE2	I feel good about myself in the process of using the e-wom platform.	
	SE3	I earn respect for myself in the process of using the e-wom platform.	
Positive emotion (PE)	PE1	I feel happy in the process of using the e-wom platform.	Fredrickson, 2013
	PE2	I feel positive in the process of using the e-wom platform.	
	PE3	I feel interested in the process of using the e-wom platform.	
Platform trust (PT)	TR1	I believe that this platform keeps its promises and commitment.	Suh and Han, 2003
	TR2	I believe that this platform meets users' expectations.	
	TR3	I believe that this platform keeps users' best interests in minds.	

confirm that respondents matched the research purposes. Finally, we collected 327 surveys, however some were discarded because respondents failed the attention check questions or replied the same answer for all items, leaving 268 valid questionnaires with a valid response rate of 81.96.

Table 2 shows the demographic information characteristics of the valid samples. Obviously, the demographic

TABLE 2 Demographic of respondents (N = 268).

Demographic variable	Frequency	Percentage
Gender		
Male	136	50.75
Female	132	49.25
Age		
18–25 years old	135	50.37
26–35 years old	123	45.90
36–45 years old	10	3.73
Education		
College and bachelor's degree	124	46.27
Master's degree	114	42.54
Doctor's degree	30	11.20
Use frequency		
Once a day or more	118	44.03
2–3 times a week	83	30.97
Once a week	11	4.10
2–3 times a month	28	10.45
Once a month or less	28	10.45

characteristics of the respondents match the users of the e-wom platform.

3.3. Data analysis

This study used partial least squares structural equation modeling (PLS-SEM) using Smart PLS 3.3 and its associated techniques, including the PLS algorithm and bootstrapping. In recent years, the number of articles published using PLS-SEM has increased significantly in contrast to covariance-based structural equation modeling (CB-SEM). Meanwhile, this study used SPSS software to perform several tests, such as descriptive and Harman's single-factor tests, which resulted in a 37.80% variation, which is less than acceptable threshold 50% (Podsakoff et al., 2003).

4. Results analysis

There are two stages to PLS-SEM analysis: measurement model and structural mode (Hair et al., 2018).

4.1. Measurement model

In this part, we accessed reliability and validity of measurement model. As shown in Table 3, regarding the reliability of the construct, the Cronbach's α of all constructs ranged between 0.830 and 0.927, which is above the acceptable threshold of 0.7 (Hair et al., 2018). The values of composite reliability (CR) ranged between 0.898 and 0.954, meeting criteria of 0.7 (Hair et al., 2018), which indicated adequate reliability. The convergent validity is assessed using two criteria, (1) standardized factor loadings of all items should exceed 0.7, and (2) the average variance extracted (AVE) of each construct needs to exceed the benchmark 0.5. As shown in Table 3, all items' factor loadings above the 0.7 threshold, and all the AVEs are above the benchmark value of 0.5. Thus, both conditions for convergent validity are adequate (Hair et al., 2018).

TABLE 3 Measurement model.

Construct	Item	Loading	Cronbach's alpha	CR	AVE
PIQ	PIQ1	0.853	0.831	0.898	0.745
	PIQ2	0.840			
	PIQ3	0.896			
PPP	PPP1	0.906	0.900	0.937	0.833
	PPP2	0.933			
	PPP3	0.899			
SSB	SSB1	0.845	0.887	0.922	0.746
	SSB2	0.880			
	SSB3	0.881			
	SSB4	0.849			
SSE	SSE1	0.939	0.927	0.954	0.873
	SSE2	0.947			
	SSE3	0.917			
PE	PE1	0.898	0.834	0.900	0.749
	PE2	0.862			
	PE3	0.836			
PT	PT1	0.882	0.830	0.898	0.746
	PT2	0.860			
	PT3	0.849			

To assess the discriminant validity, we used Fornell and Larcker's criteria (Fornell and Larcker, 1981). The square root of AVE of a construct needs to be greater than the correlation between the construct and other construct in this model. As shown in Table 4, the above criteria was clearly met.

4.2. Structural model

To assess the structural model, path coefficients, coefficient of determination (R^2), and cross-validated redundance (Q^2) were used (Hair et al., 2018). Specifically, we employed bootstrapping for the path-coefficient calculation, the PLS algorithm for the R^2 calculation, and blindfolding for the Q^2 calculation. The R^2 value evaluates the proposed model's predictive power and reflects the contribution of each construct. The R^2 value ranges between 0 and 1, where valued of 0.20,

TABLE 4 Analysis of discriminant validity.

Construct	PIQ	PPP	PT	PE	SSE	SSB
PIQ	0.863					
PPP	0.220	0.912				
PT	0.463	0.410	0.864			
PE	0.403	0.160	0.418	0.866		
SSE	0.289	0.214	0.381	0.554	0.934	
SSB	0.232	0.246	0.330	0.562	0.628	0.864

TABLE 5 Hypothesis testing and strength of the model.

Hypothesis	Path	Path coefficient	Mean	SD	*t*-value	*p*-value	Decision
Direct effect							
H1	PIQ ->SSB	0.187	0.193	0.073	2.581*	0.008	Supported
H2	PIQ ->SSE	0.255	0.258	0.070	3.646*	0.000	Supported
H3	PPP ->SSB	0.204	0.207	0.060	3.426*	0.001	Supported
H4	PPP ->SSE	0.158	0.161	0.060	2.625*	0.008	Supported
H5	SSB ->PE	0.353	0.357	0.054	6.580*	0.000	Supported
H6	SSE ->PE	0.332	0.330	0.054	6.164*	0.000	Supported
H7	PE ->PT	0.418	0.424	0.070	5.943*	0.000	Supported
Indirect effect							
H8a	PIQ ->SSB ->PE ->PT	0.028	0.030	0.015	1.837	0.065	Not supported
H8b	PIQ ->SSE ->PE ->PT	0.035	0.037	0.016	2.200*	0.026	Supported
H9a	PPP ->SSB ->PE ->PT	0.030	0.032	0.013	2.371*	0.018	Supported
H9b	PPP ->SSE ->PE ->PT	0.028	0.022	0.011	1.911	0.055	Not supported

$*t - value > 1.96 (p < 0.05)$.
R^2 (SSB) = 0.094, R^2 (SSE) = 0.107, R^2 (PE) = 0.383, and R^2 (PT) = 0.175.
Q^2 (SSB) = 0.067, Q^2 (SSE) = 0.09, Q^2 (PE) = 0.277, and Q^2 (PT) = 0.125.

0.50, and 0.75 indicate weak, moderate, and substantial effects, respectively (Hair et al., 2018). The Q^2 valued higher than zero are meaningful and that values of 0, 0.25, and 0.5 indicate small, medium, and large effects, respectively, representing the predictive accuracy of the model. This study used bootstrapping to test the path coefficients of the structural model. The results indicated that, based on the acceptance criterion (t-value > 1.96, p-value < 0.05).

The results are shown in Table 5 and Figure 2. The significant role of PIQ in driving SSB and SSE, with path coefficients of ($\beta = 0.187, t = 2.581, p = 0.010$) and ($\beta = 0.255, t = 3.646, p < 0.001$). Therefore, H1 and H2 are supported. PPP also plays a significant role in influencing SSB and SSE, with path coefficients of ($\beta = 0.204, t = 3.426, p = 0.001$) and ($\beta = 0.158, t = 2.625, p = 0.009$). Therefore, H3 and H4 are supported. SSB has a significant effect on PE ($\beta = 0.353, t = 6.580, p < 0.001$). Therefore, H5 is supported. SSE significantly affects effect on PE ($\beta = 0.332, t = 6.164, p < 0.001$). Therefore, H6 is supported. PE has a significant effect on PT ($\beta = 0.418, t = 5.943, p < 0.001$).

4.3. Mediation model

To investigate the mediating effect of SSB, SSE, and PE, the bootstrapping method was employed to estimate the indirect effect. As shown in Table 5, SSE and PE play a significant role between the relationship between PIQ and PT ($\beta = 0.035, t = 2.231, p = 0.026$); therefore, H8b is supported. SSB and PE play a significant role between the relationship between PPP and PT ($\beta = 0.030, t = 2.374, p = 0.018$). Therefore, H9a is supported. Finally, H8a ($\beta = 0.028, t = 1.837, p = 0.065$) and H9b ($\beta = 0.022, t = 1.911, p = 0.055$) are not supported. Because H8a and H9b not meet the criterion (t-value > 1.96, p-value < 0.05).

Then, in PLS-SEM, researchers generally calculate the strength of mediating effect. The variance accounted for (VAF) formula was employed (Hair et al., 2016). VAF = Indirect

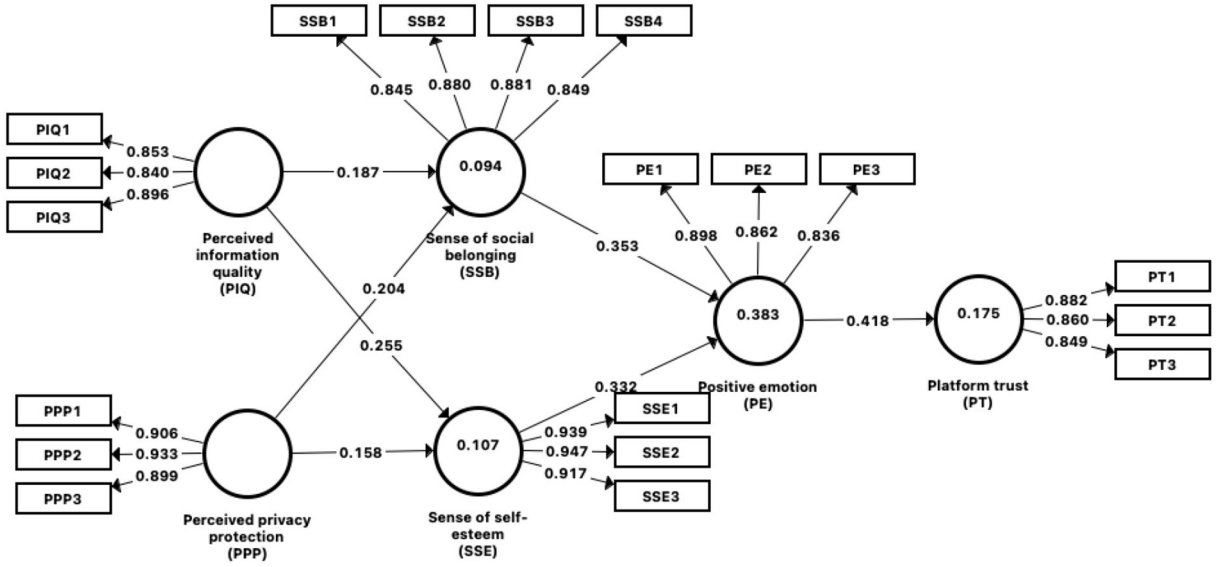


FIGURE 2
Structural model results.

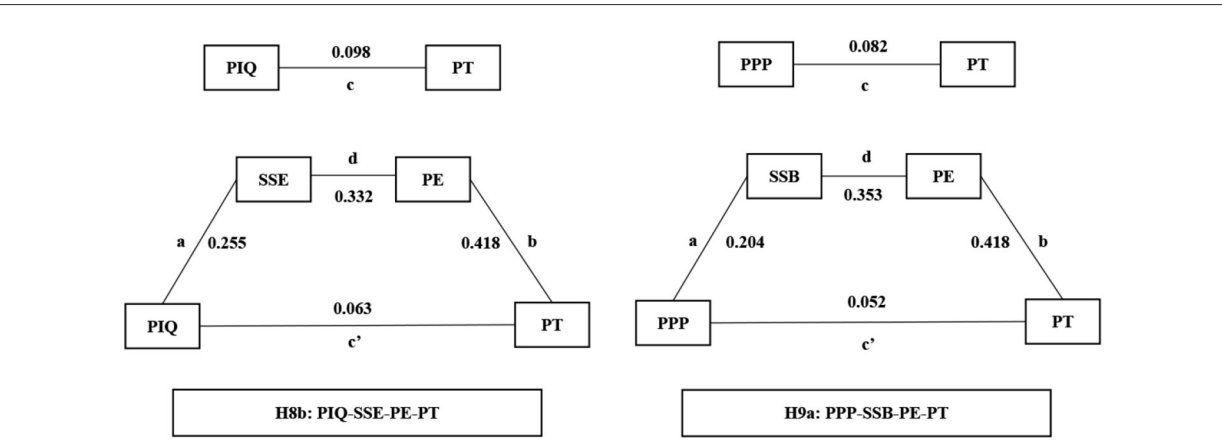


FIGURE 3
Mediation model results.

effect/Total effect, where Total effect = Indirect effect + Direct effect. VAF values of <20, 20–80, and >80% represent to no mediation, partial mediation, and full mediation, respectively (Hair et al., 2016). According to the above analysis, H8b and H9b are supported. Thus, this study calculated the strength of H8b and H9b. As shown in Figure 3, VAF values are 35.71 and 36.59%, respectively, which falls in the range between 20 and 80%, thus considered partial mediation. As shown in Figures 3, a, b, c, c′ and d represent path coefficient, a∗b∗d = indirect effect, c = total effect and c′ = direct effect.

5. Conclusion

The continued advancement of UGC and social media opens up new opportunities for both researchers and managers. The e-wom platform, as a popular information-exchange channel, requires immediate attention. Based on UG theory and S-O-R paradigm, this study investigated users' trust in the e-wom platform, especially how utilitarian gratification (perceived information quality and perceived privacy protection) affect platform trust through social gratification (sense of social belonging and sense of self-esteem) and positive emotion.

The hypotheses mentioned are supported by empirical evidence from 268 valid questionnaires. This study offers a theoretical framework for increasing users' trust in the e-wom platform.

5.1. Findings

First, positive emotion has a positive impact on platform trust. This is in line with research in the field of interpersonal trust, which shows that positive emotion helps to improve trust (Lount, 2010). When people use the e-wom platform, positive feeling displays their happiness. Users create social friendships through commenting and texting, and locate like-minded people through communities and organizations in the platform. These are all positive emotion generators. Positive emotions, such as happiness and joy, are experienced while using the e-wom platform, which increases users' trust in this platform.

Second, as previous studies shown, sense of social belonging and sense of self-esteem are determinants of positive emotion (Yagil and Medler-Liraz, 2017). This study found that social gratification (sense of social belonging and sense of self-esteem) positively influences positive emotion. In comparison to reality social circles and We-chat moments, which are made up of strong relationships, the e-wom platform is made up of weak social ties. On the e-wom platform, users are more inclined to share their inner thoughts or sentiments, show their self-image, and release the pressures of reality. Users, for example, can meet like-minded sharers and feel a sense of social connection through information on this platform. Furthermore, individuals share things that they would not dare to exhibit in public, generate specific images, and get recognition and respect from others. As a result, when users perceive sense of social belonging and self-esteem, they will be in pleasant emotional states.

Third, utilitarian gratification (perceived information quality and perceived privacy protection) positively influence social gratification (sense of social belonging and sense of self-esteem). This finding is not surprising. According to hierarchy of needs theory, personal needs progress from a low to a high condition. It follows general law of growth of personal needs to some extent. Primary function of the e-wom platform is to offer e-wom information and privacy protection. Users will have regular conversation and engagement with other users once this platform becomes an efficient information reference source. Users will show their lives, share items and services, and hint at their own taste and social position when they perceive high privacy protection. Therefore, after obtaining utilitarian gratification, users further seek for social gratification on this platform.

Finally, this study figured out the mediating effects of social gratification and positive emotion. Especially, sense of self-esteem and positive emotion play a significant role in the relationship between perceived information quality and platform trust. Sense of social belonging and positive emotion play a significant role in the relationship between perceived privacy protection and platform trust. When users perceive more information quality, sense of self-esteem and positive emotion will make them trust this platform. When this platform provides good privacy protection, it is easy for user to trust this platform once obtaining sense of social belonging and positive emotion.

5.2. Implications for research

This study contributes to social media trust research. First, prior researches primarily focused on a single component when predicting platform trust, or used platform trust as a predictor to investigate the impact on user behavior. This study investigated the dynamic mechanism of platform trust development. It expanded on the existing study framework and brought a new research perspective to the topic of diverse types of social media trust. This study expanded the relationship between self-concept and emotional states from a micro perspective.

Second, this study broadened the application of UG theory and S-O-R paradigm. This study purposed gratifications of the e-wom platform, namely utilitarian gratification and social gratification. According to S-O-R paradigm, this study treated utilitarian gratification as stimulus, social gratification and positive emotion as organism, and platform trust as response. Although there are considerable researches into users' gratification in various types of social media, few studies looked at the interaction among different types of gratifications. According to this study, different types of gratifications play different roles. Utilitarian gratification is a predictor of social gratification. This study provided a theoretical framework of hierarchical structure among types of gratifications in social media usage.

Third, we don't understand the importance of positive emotion in social media well-enough. This study investigated the impact of positive emotion on platform trust, and found that social gratification influences positive emotion. This study applied positive emotion into social media trust. This study illustrated that positive emotion plays a significant role in the domain of human-media trust.

Last but not least, this study provided a deeper understanding of the process and mechanisms that lead from utilitarian gratification to platform trust. We not only found the essential reason of e-wom platform trust, but also other factors (social gratification and positive emotion) that can contribute to it. Understanding the appearance of trust is not the purpose of this study. In-depth research of mediating effects is a key step in studying the complex variable

of trust, and it is an important part of promoting trust in human-media interaction.

5.3. Implications for practice

This study offered a new insight into how to boost trust in the e-wom platform. How to make an appealing e-wom platform that inspires users to trust it. These findings can be used to create a more persuasive and trustworthy e-wom platform. The e-wom platform provides numerous advantages for both individuals and businesses. These advantages, however, will not be achieved unless the e-wom platform establishes user trust.

First, positive emotion among users is a good predictor of platform trust. As a result, e-wom platform service providers should review user experience on a regular basis. Positive emotion is predicted by social gratification (sense of social belonging and sense of self-esteem). Service providers should organize social online activities on a regular basis to broaden users' social habits and improve user connection. Differentiating titles or interface features based on user level lets users feel valued and respected.

Then, this study discovered that utilitarian gratification is a predictor of social gratification. Therefore, service providers should improve the service quality to further promote users' social interaction within the platform. To limit the homogenization of content received by consumers under the recommendation algorithm. Content filtering technology and trusted AI technology should be used. Constantly enhancing user interface experience to improve users' sense of the effectiveness of privacy protection, such as setting up privacy policy reading links and official privacy protection push notifications.

Especially, social gratification and positive emotion have different mediating effects between utilitarian gratification and platform trust. Therefore, the platform should find the right positioning and provide different types of services for users. If the e-wom platform focuses on high-quality information, it should pay attention on providing a respectful environment. When the e-wom platform is good at protecting users' privacy, it should place particular emphasis on creating a lively social atmosphere.

5.4. Limitation and future studies

There are several limitations that need to be considered. First, because this study focuses on the e-wom platform in China, the findings may be confined to Chinese users. Future studies should increase the universality of the research model by covering a broader survey population. Second, the majority of the respondents in this study are young users. Despite the fact that the sample represents the majority of users in China, there may be differences for users of different ages. To investigate the differences in gratifications, future studies should consider using age and gender as moderators. Third, in this study, duration of collecting data is relatively concentrated. Future studies should adopt a longitudinal research approach to explore the relationships among gratification, emotion and trust.

Data availability statement

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

Ethics statement

The studies involving human participants were reviewed and approved by Beijing University of Posts and Telecommunications. The patients/participants provided their written informed consent to participate in this study.

Author contributions

XX contributed to conception, design of the study, and supervision. LL contributed to investigation, visualization, and writing—original draft. All authors contributed to manuscript revision, read, and approved the submitted version.

Conflict of interest

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

Publisher's note

References

Azer, J., and Ranaweera, C. (2022). Former customers? e-wom in social media platforms: an investigation of motives, network size and social ties. *J. Bus. Res*. 146, 118–133. doi: 10.1016/j.jbusres.2022.03.068

Bless, H., and Fiedler, K. (2006). "Mood and the regulation of information processing and behavior," in *Affect in Social Thinking and Behavior*, ed J. P. Forgas (New York, NY: Psychology Press), 65–84.

Brandtzæg, P. B., Lüders, M., and Skjetne, J. H. (2010). Too many facebook "friends"? content sharing and sociability versus the need for privacy in social network sites. *Int. J. Hum. Comput. Interact*. 26, 1006–1030. doi: 10.1080/10447318.2010.516719

Brockner, J. (1988). *Self-Esteem at Work: Research, Theory, and Practice*. New York, NY: Lexington Books.

Chai, S., Choi, B., Kim, M., and Cheng, T. C. E. (2022). Why do people speak about products online? The role of opinion leadership. *Inform. Technol. Manage*. 23. doi: 10.1007/s10799-022-00359-7

Chang, C. (2018). Understanding social networking sites continuance. *Online Inform. Rev*. 42, 989–1006. doi: 10.1108/OIR-03-2017-0088

Dunn, J. R., and Schweitzer, M. E. (2005). Feeling and believing: the influence of emotion on trust. *J. Pers. Soc. Psychol*. 88, 736–748. doi: 10.1037/0022-3514.88.5.736

Fornell, C., and Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *J. Market. Res*. 18, 39–50. doi: 10.1177/002224378101800104

Fredrickson, B. L. (2013). "Chapter 1: Positive emotions broaden and build," in *Advances in Experimental Social Psychology*, eds P. Devine and A. Plant (New York, NY: Academic Press), 1–53. doi: 10.1016/B978-0-12-407236-7.00001-2

Fredrickson, B. L., Cohn, M. A., Coffey, K. A., Pek, J., and Finkel, S. M. (2008). Open hearts build lives: positive emotions, induced through loving-kindness meditation, build consequential personal resources. *J. Pers. Soc. Psychol*. 95, 1045–1062. doi: 10.1037/a0013262

Fredrickson, B. L., and Losada, M. F. (2005). Positive affect and the complex dynamics of human flourishing. *Am. Psychol*. 60, 678–686. doi: 10.1037/0003-066X.60.7.678

Gan, C., and Li, H. (2018). Understanding the effects of gratifications on the continuance intention to use wechat in china: a perspective on uses and gratifications. *Comput. Hum. Behav*. 78, 306–315. doi: 10.1016/j.chb.2017.10.003

Goffman, E. (2016). *The Presentation of Self in Everyday Life*. Beijing: Peking University Press.

Gogan, I., Zhang, Z., and Matemba, E. (2016). Impacts of gratifications on consumers emotions and continuance use intention: an empirical study of Weibo in china. *Sustainability* 10:3162. doi: 10.3390/su10093162

Hair, J. F., Hult, G. T. M., Ringle, C., and Sarstedt, M. (2016). *A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM)*. London: Sage Publications.

Hair, J. F., Risher, J. J., Sarstedt, M., and M., R. C. (2018). When to use and how to report the results of PLS-SEM. *Eur. Bus. Rev*. 31, 2–24. doi: 10.1108/EBR-11-2018-0203

Heider, F. (1958). *The Psychology of Interpersonal Relations*. New York, NY: Wiley. doi: 10.1037/10628-000

Hsu, C.-L., and Lin, J. C.-C. (2016). Effect of perceived value and social influences on mobile app stickiness and in-app purchase intention. *Forecast. Soc. Change* 108, 42–53. doi: 10.1016/j.techfore.2016.04.012

Hsu, C.-L., and Lin, J. C.-C. (2021). The Effects of gratifications, flow and satisfaction on the usage of livestreaming services. *Library Hi Tech*. 39, doi: 10.1108/LHT-02-2021-0069

Hu, X., Chen, X., and Davision, R. M. (2019). Social support, source credibility, social influence, and impulsive purchase behavior in social commerce. *Int. J. Electron. Commerce* 23, 297–232. doi: 10.1080/10864415.2019.1619905

Huang, L., Wu, N., You, Z., Liu, G., and Zhou, Z. (2016). Emotional experiences and users' behaviors in social networking services - an study of sina microblog. *China J. Clin. Psychol*. 24, 409–412. doi: 10.16128/j.cnki.1005-3611.2016.03.006

Katz, E., Blumler, J. G., and Gurevitch, M. (1974). "The uses of mass communications: current perspectives on gratifications research," in *Utilization of Mass Communication by the Individual*, eds J. G. Blumler and E. Katz (Los Angeles, CA: Sage Publications), 19–13.

Kim, D. J., Ferrin, D. L., and Rao, H. R. (2008). A trust-based consumer decision-making model in electronic commerce: the role of trust, perceived risk, and their antecedents. *Decis. Support Syst*. 44, 544–564. doi: 10.1016/j.dss.2007.07.001

Kim, H. (2021). Use of mobile grocery shopping application: motivation and decision-making process among South Korean consumers. *J. Theor. Appl. Electron. Commerce Res*. 16, 2672–2693. doi: 10.3390/jtaer16070147

Lavy, S., and Littman-Ovadia, H. (2017). My better self: Using strengths at work and work productivity, organizational citizenship behavior, and satisfaction. *J. Career Dev*. 44, 95–109. doi: 10.1177/0894845316634056

Leung, L., and Wei, R. (2000). More than just talk on the move: uses and gratifications of the cellular phone. *J. Mass Commun. Q*. 77, 308–320. doi: 10.1177/107769900007700206

Li, W., and Lin, S. (2021). Research on trust repair mechanism after social media privacy invasion. *Library Inform. Service* 65, 33–44. doi: 10.13266/j.issn.0252-3116.2021.17.004

Lin, H.-F. (2008). Determinants of successful virtual communities: contributions from system characteristics and social factors. *Inform. Manage*. 45, 522–527. doi: 10.1016/j.im.2008.08.002

Lin, J., Guo, J., Turel, O., and Liu, S. (2020). Purchasing organic food with social commerce: an integrated food-technology consumption values perspective. *Int. J. Inform. Manage*. 51:102033. doi: 10.1016/j.ijinfomgt.2019.11.001

Liu, Z., Qiao, Y., and Wu, R. (2018). Predicting mechanism of self-esteem on meaning in life: a dual mediation model of locus of control and positive emotion. *Chinese Ment. Health J*. 32, 772–777. doi: 10.3969/j.issn.1000-6729.2018.09.012

Livingstone, S. (2008). Taking risky opportunities in youthful content creation: teenagers? use of social networking sites for intimacy, privacy and self-expression. *New Media Soc*. 10, 393–411. doi: 10.1177/1461444808089415

Lount, R. B. (2010). The impact of positive mood on trust in interpersonal and intergroup interactions. *J. Pers. Soc. Psychol*. 98, 420–433. doi: 10.1037/a0017344

Lu, X., Phang, C. W., and Yu, J. (2011). Encouraging participation in virtual communities through usability and sociability development. *ACM SIGMIS Database* 42, 96–114. doi: 10.1145/2038056.2038062

Luo, B., and Cong, R. (2015). Retention, transmission, search and use: review and prospect of e-word-of-mouth research from the perspective of consumer behavior. *For. Econ. Manage*. 37, 54–64. doi: 10.16538/j.cnki.fem.2015.08.007

Meng, Z. (1989). *Human Emotion*. Shanghai: Shanghai People's Publishing House.

Molinillo,S., Aguilar-Illescas, R., Anaya-Sanchez, R., and Liebana-Cabanillas, F. (2021). Social commerce website design, perceived value and loyalty behavior intentions: the moderating roles of gender, age and frequency of use. *J. Retail. Consum. Services* 63:102404. doi: 10.1016/j.jretconser.2020.102404

Pan, X., Qin, Q., Zhang, Y., and Tan, X. (2012). The effect of organizational psychological ownership and organization-based self-esteem on positive organizational behaviours. *J. Psychol. Sci*. 35, 718–724. doi: 10.16719/j.cnki.1671-6981.2012.03.034

Park, M.-S., Shin, J.-K., and Ju, Y. (2019). Attachment styles and electronic word of mouth (e-wom) adoption on social networking sites. *J. Bus. Res*. 99, 398–404. doi: 10.1016/j.jbusres.2017.09.020

Podsakoff, P. M., Mackenzie, S. B., Lee, J.-Y., and Podsakoff, N. P. (2003). Common method biases in behavioral research: a critical review of the literature and recommended remedies. *J. Appl. Psychol*. 88, 879–903. doi: 10.1037/0021-9010.88.5.879

Rosenberg, M. (1965). *Society and the Adolescent Self-Image*. Princeton, NJ: Princeton University Press. doi: 10.1515/9781400876136

Rousseau, D. M., Sitkin, S. B., Burt, R. S., and Camerer, C. (1998). Not so different after all: a cross-discipline view of trust. *Acad. Manage. Rev*. 23, 393–404. doi: 10.5465/amr.1998.926617

Sapuppo, A., and Seet, B.-C. (2015). Privacy and technology challenges for ubiquitous social networking. *Int. J. Ad Hoc Ubiquit. Comput*. 18, 121–138. doi: 10.1504/IJAHUC.2015.068127

Schwartz, B. (1968). The social psychology of privacy. *Am. J. Sociol*. 73, 741–752. doi: 10.1086/224567

Schwarz, N., and Clore, G. L. (1983). Mood, misattribution, and judgments of well-being: informative and directive functions of affective states. *J. Pers. Soc. Psychol*. 45, 513–523. doi: 10.1037/0022-3514.45.3.513

Suh, B., and Han, I. (2003). The impact of customer trust and perception of security control on the acceptance of electronic commerce.

Int. J. Electron. Commerce 7, 135–161. doi: 10.1080/10864415.2003.11044270

Wang, H., and Lian, X. (1998). A dyadic perspective: the research on the dynamicly interactive relationship between trustor - trustee. *J. South China Univ. Technol.* 18, 27–35.

Yagil, D., and Medler-Liraz, H. (2017). Personally committed to emotional labor: Surface acting, emotional exhaustion and performance among service employees with a strong need to belong. *J. Occup. Health* 22, 481–491. doi: 10.1037/ocp0000049

Yuan, B., Sun, X., You, R., Liu, F., and Li, W. (2018). The effect of emotion on trust: a systematic review and meta-analysis. *Stud. Psychol. Behav.* 16, 632–643. doi: 10.3969/j.issn.1672-0628.2018.05.008

Yuan, S., Liu, L., Su, B., and Zhang, H. (2020). Determining the antecedents of mobile payment loyalty: cognitive and affective perspectives. *Electron. Commerce Res. Appl.* 41:100971. doi: 10.1016/j.elerap.2020.100971

Zheng, X., Li, Y., and Liu, Z. (2017). The impact of knowledge sharing on employees innovative behavior: the role of organization based self-esteem and perceived organizational support. *J. Bus. Econ.* 37, 24–33. doi: 10.14134/j.cnki.cn33-1336/f.2017.01.003

frontiers | Frontiers in Computer Science

TYPE Original Research
PUBLISHED 16 September 2022
DOI 10.3389/fcomp.2022.936146

OPEN ACCESS

EDITED BY
Jason Nurse,
University of Kent, United Kingdom

REVIEWED BY
Evellin Cardoso,
Universidade Federal de Goiás, Brazil
Muhanna Muhanna,
Princess Sumaya University for Technology, Jordan

*CORRESPONDENCE
Lucia Vilela Leite Filgueiras
lfilguei@usp.br

SPECIALTY SECTION
This article was submitted to Human-Media Interaction, a section of the journal Frontiers in Computer Science

RECEIVED 04 May 2022
ACCEPTED 12 August 2022
PUBLISHED 16 September 2022

CITATION
Filgueiras LVL, Corrêa PLP, Alves-Souza SN, Teodoro SM, Silva MSP, Encinas Quille RV and Demuner VRS (2022) Working with robotic process automation: User experience after 18 months of adoption. *Front. Comput. Sci.* 4:936146. doi: 10.3389/fcomp.2022.936146

Working with robotic process automation: User experience after 18 months of adoption

Lucia Vilela Leite Filgueiras[1]*, Pedro Luiz Pizzigatti Corrêa[1,2], Solange N. Alves-Souza[1], Sigmar Monroe Teodoro[3], Mariana Savarezze Pereira da Silva[3], Rosa Virginia Encinas Quille[1,2] and Vanessa Rafaela de Souza Demuner[3]

[1]Polytechnic School, University of São Paulo, São Paulo, Brazil, [2]School of Arts, Sciences and Humanities, University of São Paulo, São Paulo, Brazil, [3]EdP Brasil, São Paulo, Brazil

This paper reports a study of User Experience (UX) with Robotic Process Automation (RPA), in the perspective of workers of EdP Brazil, a large electric utility company that operates in Brazil. RPA are software solutions for automating business processes that find increased interest of companies because they are inserted in workgroups as a co-worker, emulating human workers operating on GUI interfaces. Although the technology promises to drive a new wave of productivity in service companies, its impact on co-workers' experience is still unexplored. Based on projective interviews using the AXE (Anticipated eXperience Evaluation) protocol, after the first 18 months of RPA operation, the analysis of workers' collaboration with the robots has evidenced multiple facets of UX, technology acceptance and innovation adoption. For this case, RPA has provided an overall positive user experience mainly due to the perceived utility of the spared time, the upgrade in career opportunities and the pride for actively participating in the innovation adoption. Negative experience comes mainly from the lack of visibility that hinders robot management for efficiency and improvement. The methodology used in the study was successful in capturing the multifaceted workers' experience and is potentially useful to support user research in new expansion RPA projects.

KEYWORDS

acceptance models, human-robot interaction, innovation adoption, robotic process automation, user experience (UX), human-computer interaction

1. Introduction

Robotic Process Automation (RPA) are software agents that automate clerical manual tasks by processing data. While robotic process automation technology has several, clearly defined benefits for the company, workers' experience with the robots are still not well documented in the literature. The IEEE Guide for Terms and Concepts in Intelligent Process Automation defines RPA as: "A preconfigured software instance that uses business rules and predefined activity choreography to complete the autonomous execution of a combination of processes, activities, transactions, and tasks in one or more unrelated software systems to deliver a result or service with human exception management" (Group, 2017).

RPA software, from now on designated simply as RPA, usually operate business information systems as a person would do, by retrieving data from business applications, filling forms and processing transactions under business rules (Alberth and Mattern, 2017). RPA software is configured to connect to an ERP system and/or other corporative systems, from which it extracts business data, replacing humans in well-defined manual and repetitive tasks.

As Wilcocks and colleagues state, "RPA software is ideally suited to replace humans for so called "swivel chair" processes; processes where humans take inputs from one set of systems (for example email), process those inputs using rules, and then enter the outputs into systems of record (for example Enterprise Resource Planning (ERP) systems)" (Willcocks et al., 2015).

RPAs can be grouped into three classes, or "generations": the first one, labeled G1 RPA, automate tasks based on structured data found on systems databases. The second class, labeled G2 RPA, works from unstructured data such as text files based on machine learning techniques. G3 RPAs are cognitive platforms that can perform decision-making tasks (Ernst and Young, 2015).

G1 RPA is an interesting solution to lessen operational costs, as its adoption reduces operational costs with little impact on IT infrastructure. Robots interface with existing systems by emulating a human user, and they can be configured by business rule experts instead of IT personnel. The potential to use cognitive solutions in G2 and G3 RPA promises to increase transformation in the scenario of process automation in the next few years (Rutaganda et al., 2017). Effectively, according to MarketsandMarkets, the market for RPAs was estimated to grow over 30% between 2017 and 2022, mainly driven by "the ease of business processes provided by the robotic process automation, and convergence of robotic process automation with traditional business process industries" (MarketsandMarkets, 2017).

Claims toward adopting RPAs are based on several arguments: RPA potentially reduces FTE (full-time employees); RPA provides increased service quality due to the fact that robots do not make mistakes while doing what they are programmed to do: RPA increases efficiency due to the robots overall performance, superior to human's; RPA increases liability, because the transactions can be automatically documented according to compliance requirements; RPA provides uplifting of the human workforce, because being free of repetitive, tedious work, people can perform more valuable tasks such as listening to customers, analyzing the business and innovating.

Following the international trend EdP Brazil has started adopting G1 RPA in 2017 as part of the company effort to improve efficiency and reduce the risk of costs associated to non-compliance with the national energy agency rules. EdP Brazil implemented RPAs with Blue Prism™ technology in several business units, mainly in the financial areas. EdP Brazil is progressing in this process by developing its first G3 RPA (Vajgel et al., 2021).

Additionally, EdP led the foundation of the Brazilian Business Pact for Humanized Work Digitalization, which established principles to promote a human-centered process for robotization and digital transformation. These principles focus on empowering humans by education, inclusion, engagement, leadership development and compliance with ethical behavior. As a result of this pact, the authors, researchers from Universidade de São Paulo (USP) in Brazil and EdP Brazil employees in the Research, Development and Innovation area have cooperated to study workers' experience with RPA, under PROP&D, a R&D program regulated by the Brazilian Agency for Energy, ANEEL

In fact, while robotic process automation technology has several proven benefits for the company, user experience is a strategic factor for technology acceptance that is still poorly documented in the scientific literature.

Recent systematic literature reviews show that there is a growing interest in the area (Ivančić et al., 2019) as RPA is being massively applied to industry. A literature review by Syed et al. (2020) has pointed to several relevant research challenges. Yet, the user experience (UX) perspective of technology has not deserved a dedicated look in academic studies. Our own literature review showed that most of the publications about RPA belong to the business administration area, reporting on the success factors of RPA in business (Lacity et al., 2016; Rutaganda et al., 2017; Devarajan, 2018; Leshob et al., 2018) and on applications in different areas such as healthcare (Ratia et al., 2018), public administration (Houy et al., 2019), software engineering (Montero et al., 2019), and finance (Stople et al., 2017; Lewicki et al., 2019). Others are focused on identifying which tasks are suitable for being automated by RPAs (Vishnu et al., 2017). However, the effectiveness of innovation in a company depends not only on the technology, but also on factors that express how workers understand changes and cope with them.

Very few studies were found that considered the human factor with RPA in depth. Our literature search for related studies addressed the area of Human-Robot Interaction (HRI), which studies the interaction between robots and humans (Fontanillo Lopez et al., 2020; Schellen et al., 2021). Although the subject of relationship between humans and automata is not new, the study of HRI became more prominent in the 1980s, when behavior-based robotics started using distributed sense-response loops to generate appropriate responses to external stimuli. According to Goodrich and Schultz in their review on HRI, an HRI problem consists in understanding and shaping the interactions between one or more humans and one or more robots (Goodrich and Schultz, 2008). Chen and Barnes have reviewed the literature on human-agent interaction to identify the most critical issues that need to be addressed for such systems to be effective (Chen and Barnes, 2014). Yet although their discussion is applicable to RPA and RPA is a legitimate HRI problem,

human-RPA interaction has not been studied yet in the HRI domain.

In the Computer-Supported Cooperative Work area, we found that Nauwerck and Cajander have published their preliminary results of a study of human introduction of robotic process automation (RPA) of financial support in a Swedish municipality focusing on workers' experience (Nauwerck and Cajander, 2019). However, the most comprehensive study found about workers' relationship with RPA was Katriina Juntunen's thesis from the Aalto University RPA adoption and acceptance processes in the Financial department of Stora Enso, a Finnish paper manufacturer (Juntunen, 2018). Her work produced an integrative framework in which she analyzed the organizational factors that influenced the adoption of RPA. This framework was helpful for our analysis in this paper, as we present in the following sections.

We have here attempted to understand how EdP workers relate to G1 RPAs after 18 months of its adoption. Our research question is stated as "After 18 months of the decision to adopt RPA in EdP, how do workers characterize their experience with RPA technology?" This research question is decomposed into three secondary questions:

RQ1: Which factors have influenced their positive experience with RPA?
RQ1: Which factors have influenced their negative experience with RPA?
RQ1: Which are the expectations of these workers regarding this technology evolution?

This paper is organized in 7 sections. In section 2, we present the theoretical background that supports this investigation. In Section 3, we describe the scenario of RPA application in EdP. In Section 4, we detail our research methodology. Section 5 contains the presentation and discussion of our findings regarding the UX framework. Section 6 presents our findings regarding the acceptance and adoption framework. Section 7 presents our discussion on findings as well as our reflections on how to improve workers' experience with RPAs.

2. Theoretical background

This research is based on theories and methods of the Human-Computer Interaction area, in which phenomena related to how people interact with technology are the core interest. At first, we have based our research on User eXperience (UX) frameworks, but the preliminary results suggested an expansion of analysis to broader organizational frameworks. In this section, we present the frameworks that guided our study.

2.1. User experience

Understanding UX is essential in modern design approaches. In the Human-Centred Design perspective, the design of an innovation must keep a close focus on humans, their needs and characteristics (FDIs, 2008).

Among the many definitions of UX, we pick Hassenzahl and Tractinsky's for its coherence with our analysis target: user experience is a consequence of a user's internal state (predispositions, expectations, needs, motivation, mood, etc.), the characteristics of the designed system (e.g., complexity, purpose, usability, functionality, etc.) and the context (or the environment) within which the interaction occurs (e.g., organizational/social setting, meaningfulness of the activity, voluntariness of use, etc.) (Hassenzahl and Tractinsky, 2006).

UX is unique to an individual. UX definition restrains the experience to those encountered with some technological artifact, be it a system, a product or a service. It is the result of the individual's encounters with the artifact, and is affected by prior experiences, expectations, cultural background and social context. UX refers to both the cognitive and emotional consequences of encounters.

As encounters can happen over time, UX varies due to exposure to systems. Roto, Law, Vermeeren and Hoonhout, early researchers on the concept of UX, have identified time spans of UX. People can have anticipated UX before their first encounter with the target technology, based on expectations, experience with other systems or information about the new technology. Momentary UX is related with one single encounter and to the feelings brought by that instantaneous interaction event. We refer to episodic UX when we consider a certain usage situation that may have happened in the past. Over time, the series of momentary experiences result in cumulative UX, which is defined as the reflection over the recollection of various episodes of usage. Interestingly, anticipated UX may relate to cumulative UX, because expectations are constructed based on previous experience (Roto et al., 2011b).

UX is influenced by factors, namely system, user and contextual factors (Roto et al., 2011b). Different UX authors identify different attributes for these factors. System factors are related to usability and quality-in-use attributes, as defined in ISO SQuaRE system (ISO, 2011). User factors may include prior knowledge and willingness to use the system as well as affective characteristics and personality traits. Contextual factors of different natures, from social and organizational environment to physical conditions and task specifications, can also shape experience.

The complexity of multiple influential factors added to the individual nature of UX makes experience design a difficult undertaking. However, it is possible to design for UX (Hassenzahl, 2013). Understanding influential factors is strategic knowledge; currently, companies are largely investing in UX to

identify their customers' values, to propose new products, to make processes more efficient and to reduce waste.

There are several approaches to UX. Among different models and methods, Hassenzahl's framework for understanding UX is widely recognized as a comprehensive model, encompassing factors of different natures (Hassenzahl, 2003). Hassenzahl's model explains that a product is designed to present an Intended Quality. Intended qualities can be both pragmatic, that is, associated to what the system can do in both the functional and non-functional perspectives, and hedonic, associated to the emotional consequences that designers intend their users to experience. In real use, the designed intended quality is somehow perceived by users. The perceived pragmatic quality and the perceived hedonic quality can be different from the designed ones. They are compiled by a user's cognitive assessment into product attractiveness, which results in behavioral consequences (such as increased use) and emotional consequences (such as joy).

2.2. Methods for UX evaluation

Academic and professional literature also show several techniques and tools to evaluate UX. The website All about UX (https://www.allaboutux.org/all-methods, accessed on 07/30/2022) lists 81 different methods and techniques for evaluating UX. In this research, we chose the AXE-Anticipated eXperience Evaluation method because of its support to a qualitative in-depth interview, to the problem we had in hands. The AXE protocol is briefly described as follows.

AXE is a qualitative user research method developed by Lutz Gegner and Michael Runonen of Aalto University, in a cooperation project between Departments of Design and of Computer Science of Aalto University and Nokia Research Center, Helsinki. The protocol was developed for evaluating interactive concepts and early prototypes under the authors' assumption that "identifying important experiential aspects during very early phases of development can reduce costly changes but also provide a competitive edge." (Gegner and Runonen, 2012).

The AXE protocol is based on psychological projective tests. Participants are shown pictures they associate with the product, system or service that is the target of evaluation. The pictures, predefined in the protocol, were selected to evoke the concepts of hedonic and pragmatic qualities, as well as attractiveness, from Hassenzahl's user experience model in the AttrakDiff questionnaire (Hassenzahl, 2003). Pictures are deliberately ambiguous so that participants can interpret them according to their personal background and beliefs, and express their "attitudes, opinions and self-concept" about the target product. As the AXE authors advocate, the activity of freely interpreting a picture helps remove the interference of the interviewer's words with the recall of the participant's experience.

Despite conceived as a method for evaluating experiences with early prototypes, our previous experience with AXE indicated that it is an interesting method also for cumulative UX, when the user recollects multiple periods of use (Roto et al., 2011a). The projective characteristic of the interview allows the moderator to place the focus on any moment in the experience timeline.

Also, we predicted that user experience with a new technology in a work environment with many context variables could have strong influence from organizational factors not captured by UX assessment methods with structured interviews, such as AttrakDiff. We understood that in-depth interviews could expose the totality of the experience. In previous studies with the AXE method, we observed that participants' interpretation of the images stimulated them to talk about what is relevant to them, even if not directly related to the hedonic and pragmatic qualities of the model underpinning the method.

AXE framework for analysis is based on Hassenzahl's UX model (Hassenzahl, 2003). There are three main categories, Perceived Product Features, Associated Attributes and (Anticipated) Consequences.

According to the protocol handbook (Gegner and Runonen, 2012), the analyst must classify under the Perceived Product Features category users' appreciation of the system look&feel, that is, opinions on Content, Interaction, Presentation and Functionality features.

The *Associated Attributes* category must be used to compile users' appreciation of system qualities. Opinions must be further separated into attributes associated to the system meeting *Pragmatic Needs* (*Utility* and *Usability*) and attributes associated to the system meeting users' Hedonic Needs (*Stimulation, Identification and Evocation*).

Reports on how the user felt attracted or changed their behavior are classified under Attractiveness and Behavioral Change subcategories of the (*Anticipated*) Consequences category.

Users may also express their perceptions as Suggestions or criticisms (*Unwanted*). Opinions about the overall concept are grouped into a Meta category.

2.3. Tecnhology acceptance and innovation adoption models

The trend toward RPA technology in companies has motivated researchers to study this technology from the point of view of technology acceptance models, as well as innovation adoption models.

Technology acceptance models intend to explain the user's decision about using or not a given technology. These frameworks focus on understanding the motivation of an individual to a certain behavior thus explaining the adoption

of technology from an individual basis. Innovation adoption models propose a sequence of steps that an organization should go through in order to decide whether to adopt or reject a technology. They study how an idea perceived as new spreads through a social system and gets incorporated. Both kinds of models have constructs that help explain the relationship between users and an innovative technology. They were proven helpful to show the reasons why a technology may succeed or fail in their real application.

There are many contact points between UX models and technology acceptance models. Factors that provide a positive user experience can influence the acceptance behavior. Used in combination, the constructs of these models can be effective for understanding people's intention to use a technology (Prietch and Filgueiras, 2015; Al-Rahmi et al., 2019).

Our literature research on RPA technology adoption and acceptance resulted in one relevant publication, a Masters dissertation by Katriina Juntunen from Aalto University (Juntunen, 2018), in which she analyzed the intra-organizational adoption of RPAs in the Financial department of Stora Enso, a Finnish paper manufacturer.

In that work, Juntunen analyzed and compiled constructs from 8 models that explain adoption from the individual, social and managerial perspective: four user acceptance models, including the well-known Technology Acceptance Model (TAM), by Davis and colleagues (Davis, 1989) and the Unified Theory of Acceptance and Use of Technology (UTAUT), by Venkatesh and collaborators (Venkatesh et al., 2012); one innovation diffusion model, the Innovation Diffusion Theory (IDT) by Rogers (Rogers et al., 2014); and three change management models, which were included to provide a clear view on how management acts to promote innovation adoption.

Juntunen's composite model resulted in a structure for understanding the intra-organizational adoption of innovation. She proposes that influencing variables and key *beliefs* condition adoption.

Influencing variables, summarized in Table 1, were derived from all the cited models and include innovation attributes, organizational attributes, individual attributes and managerial facilitation attributes. In section V herein, we present the definition of each variable together with our interpretation of this variable in the context of our study.

The model also establishes key beliefs. The concept of belief is not clear in the literature and Juntunen does not define her position; however, based on the discussion by Österholm (2010), we assume that belief is the knowledge a person assumes to be true in the context of their actions, that is, the attitude toward innovation depends on the beliefs about the consequences of performing the behavior and evaluation of these consequences. She defines five influential beliefs that define whether individuals accept or resist the change, so that these beliefs should be influenced to promote the change:

TABLE 1 Juntunen's composite model of technology acceptance and adoption-influencing variables (Juntunen, 2018).

Category	Attribute
Innovation attributes	Relative advantage
	Complexity
	Trialability
	Observability
	Job-fit
	Voluntariness
Organizational attributes	Compatibility
	Organizational norms
	Innovativeness
	Resource factors
	Use and support of others
Individual attributes	Personality
	Socio-economic factors
	Communication behavior
	Innovativeness
	Gender
	Age
	Expertise
Managerial facilitation	Active participation
	Human resources management
	Management of information
	Persuasive communication
	Formalization activities
	Diffusion practices
	Rites and ceremonies

- Perceived benefits, related to individuals' perception of outcomes and benefits of behavioral change.
- Perceived effort, related to individuals' perception of the number of resources to be dedicated to the change.
- Perceived social pressure and influence, that relates to the risk of being against the social tendency.
- Perceived need and appropriateness, that express the understanding of personal and task needs and the appropriateness of innovation.
- Perceived capabilities, which address how the individual evaluates self-efficacy and capability to perform, as well as resources availability.

Juntunen's framework complemented the UX framework as it explains several constructs that can influence experience but are encapsulated as "contextual factors." In turn, as the paper reveals, the UX model complemented the adoption framework by explaining the human reasons behind constructs.

3. Scenario of RPA application

In this section, we describe the scenario in which we carried out our analysis. EdP Brasil has been consolidating its position as one of the largest companies in the electric energy area in Brazil, through strategic acquisitions of energy companies in the sector's privatization process. With this, the company inherited from the acquired companies business processes, teams, and systems, which need to go through revision and standardization. These activities are part of the company's digital transformation process, and in this context, the adoption of RPAs is seen as a necessary and urgent transformation.

When the introduction of RPA was initiated, business units were advised regarding the required characteristics of processes to be automated, considering the level and capabilities of the G1 RPA available to the organization. Automation candidate activities were those that fully matched requirements of data volume and execution time and presented a standardized execution process. Although the company had established processes at the value chain level, operational processes often lacked standardization and documentation, in which each employee performed the activity differently. Thus, before automation could proceed, business units was required to apply well-known strategies and models such as PDCA, Lean and 5W2H to standardize their candidate processes to automation. Our research was performed in the context of the client support activity in EdP Brazil which had gone through this effort. Two independent business processes automated with RPA technology were selected for analysis: the damage compensation process and the billing anomaly process. We describe them briefly in the following sections.

3.1. Damage compensation process

The damage compensation process (internally known as PID, from its Portuguese abbreviation) is responsible for compensating for damages caused by energy fluctuation to customers' electrical equipment. Electrical fluctuation can happen in the distribution network due to several causes, such as storms and equipment failure. In summer rainy seasons, the number of complaints due to electrical damage rises significantly. The company must respond to complaints in due time, as defined in Chapter XVI of the Brazilian ANEEL Regulatory Resolution 414/2010. Ṙobots are employed to scrutinize databases for evidence of matching between electrical incidents, affected areas and the customers' reports. Also, they organize communication with clients and follow the document exchanges for compensation, thus avoiding fines and economic penalties that result when the deadline for analysis and response to clients is not met.

3.2. Billing anomalies process

The second process deals with anomalies in billing accounting. Each day, a list of non-conformities in payments is detected by the billing system. Non-conformity causes are various; they can be divergence between values, error in barcode typing, errors in bank reports, wrong values, to cite a few. These cases are named anomalies. Robots diagnose and solve anomalies caused by known situations, performing analysis on several parameters and checking them in different systems. When the robot succeeds in classifying the anomaly in one of the known cases, it is corrected and cleared. If the robot is not capable of identifying the cause of anomaly, it reverts to manual.

4. Methodology

Our study was a qualitative research, guided by UX and technology adoption frameworks. We collected data on UX using AXE protocol with EdP employees, who are related to the robotized processes described in Section 3. Subjects were recruited by the company R&D managers based on their experience with the RPA technology. They are real users of the robot or process managers. Some of them participated in the robot configuration and deployment.

The AXE application has three phases: preparation, interview and analysis. In preparation phase, we provided two additional pairs of pictures that address specific features of the product as described in Section 4.2. Interviews took place in meeting rooms isolated from the sight and sound of the workplace, with only the participant and the interviewer (the first author of this paper) side by side at the table. Personnel and equipment were kept to a minimum: the interviewer, her laptop running a form for participant identification and her smartphone running an audio recording app.

The evaluation session was carefully explained in the context of the R&D project. Emphasis was given to the anonymization process of results in which the researcher removed names and text excerpts that could potentially identify the subject. Participants were informed that they were free to refuse to participate and invited to sign a consent form in case they agreed.

The submission of the research project to an Institutional Review Board was not required under the Brazilian regulation for ethics in research involving human participants (Brasil et al., 2016). This regulation exempts from submission and approval research projects that aim at the theoretical deepening of situations that emerge spontaneously from professional practice, provided they do not disclose data that can identify the subject.

Since the business process is well-known by the users, they were invited to recall a normal workday without the robot and a normal workday with the robot. Next, a welcome page was presented and read together. The second page had

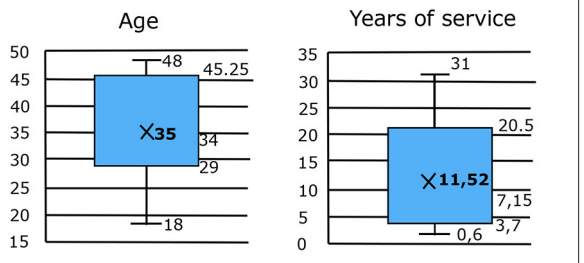


FIGURE 1
Participants' age and years of service in EdP.

the instructions and the warm-up pair. The warm-up pair is easily interpreted as a concept of speed and performance. The interviewer presented two different interpretations of the picture to show that there is no right or wrong answer.

The pairs of pictures were then presented, and the interviewer prompted the participants to explain their choices and to discuss important matters as the interview proceeded. The interviews were informal and relaxed. Interviews were transcribed and temporized.

Segments were selected; snippets were coded using Grounded Theory for concepts (Corbin and Strauss, 2008), together with codes from both AXE and Juntunen's frameworks using Qualitative Data Analysis software.

4.1. Participants

We interviewed 10 workers of EdP Brazil, five males and five females. Five participants work in line with the robots, that is, their work activity requires continued contact with the robot. Two participants work occasionally in line with the robot, that is, they have other attributions, but they may be asked to work in the automated process when the workload is excessive. Three participants are managers: two are the owners of the selected processes and one is the manager of the automation robot initiative.

Three participants have finished graduate studies, six have finished college and one is a college student, working as intern in the team. College titles are varied, from humanities to engineering. Figure 1 shows the age distribution and years of service in the company.

4.2. Pairs of pictures

Choice of image pairs followed the guidelines for selecting pictures in AXE handbook (Gegner and Runonen, 2012). In these guidelines, AXE authors recommend choosing pairs of images with at least two easily identifiable opposite attributes. Also, they advise that some ambiguity is preserved, so that they provide a lead for discussion. Additional image pairs should motivate users to talk about aspects of interest to the application project. Our research team had identified, in discussions leading up to the project, concerns from RPA managers about workers' experience with RPA being an obscure technology, and that it caused social tension by the prospect of unemployment. Pictures were picked from internet image files. The first pair of pictures was chosen to convey the opposition between company and individual; rich and poor; employee and customer. Thus, the left side showed a young man working at his computer while the right side showed houses in a low-income district. The second pair presented the surface of a moon full of craters in a dark space, in opposition to a stream of clear water being poured into a metal spoon. This pair intended to provoke the opposition between difficult and easy to learn; magical/obscure and explicit; unknown and known.

4.3. Text segmentation and coding

The transcript of interviews was segmented into 1,050 snippets, which were then coded using AXE and Juntunen's frameworks. The analysis of these data produced 100 facts, which are discussed in the following sections.

5. Results of UX evaluation

In this section, we present findings from the qualitative data analysis relative to UX. We used the AXE framework to disclose workers experience with RPA. Each finding is explained and evidenced by excerpts of interviews. The participants are referred to as (Si).

Statements are grouped by UX attributes, and those are grouped into categories according to the analysis framework.

AXE categories were used to group our findings. Since the experience totalizes several influential factors, often one experiential expression holds elements associated to more than one category. Our intention of grouping into categories only intends to organize the presentation of findings, and not to trace any solid line between categories. Table 2 summarizes the findings, detailed in the following sections.

5.1. Perceived features

This category is related to the workers' perception on the robot's characteristics as an interactive system.

RPA is not a canonical interactive software—robots have no clear user interface that the user can manipulate and from which the user obtains information. The interface between robot and users is a file containing the identifiers of the

TABLE 2 User experience with RPA.

Category	Attribute	Findings
Perceived features		
	Content	Process organization
		Activities standardization
		Opacity of complex process
		Workers as robots' redundancy
		Unpredictable workload
		Missing indicators of robots' performance
	Interaction	Humans as robots' supervisors
	Functionality	Robots not allowed to fully perform their functionality
		Robots underused and sub-effective
	Presentation	Prompt response
		Lack of status visibility
		Minimalist communication
		Supervision by indirect evidences
		Lack of information on error
Associated attributes		
	Evocation	Usefulness of released work time
		Tranquility
		Engagement
		Robots as trustworthy colleagues
		Positive surprise with effectiveness
		Loneliness
		Expectation of unlimited possibilities of growth
	Identification	Professional reward from robots' success
		Responsibility for new careers
		Maturity growth
		Human still smarter than robots
		Pride for successful implementation
	Stimulation	Change from tedious work
		Increased influence due to deeper understanding of the company's business process
		Work experience as employment advantage
		Frustration when robots fail
	Usability	Robots free from errors
		Fast robots
		Effective and resource efficient robots
		Accurate robots
		Improved overall process quality
	Usefulness	Robots as useful teammates
		Utility compromised by business changes.
		Robot improvement dependent on managerial decision to invest.
Anticipated consequences		
	Attractiveness	Broadening scope of robotization
		Non-digital or non-integrated sources of relevant data
		Intelligent robots
		Robots to resemble humans
		Concern about job loss in the future
		Workers' innocence of the unemployment process
	Behavior changes	Change toward attitude of creative restlessness
		Increase in interpersonal distance
		People growth due to innovation
Suggestions		
		Broadening of robots' scope
		Management of robots' performance on operational team's hands
		Frequent revisions of operational procedures.

transactions that must be executed by the robot, in chronological order. As the robot successfully performs the transaction, the corresponding register is removed from the list. Workers then observe the decrease in the number of records to assess robot operation. If an operation fails, the register is marked with an icon symbolizing that the transaction has been processed but terminated unsuccessfully. In the latter case, the user retrieves the identifier to analyze the reason for non-execution and, when possible, executes the activity manually. Usually, the reason is the failure of an application or a change in structure/layout of a site which the RPA accessed to perform some query. Also, the user receives a daily email with the list of transactions that are exceptions to the standardized process and that should be processed manually. Although rudimentary, there is interaction in the collaboration between humans and robots and, unquestionably, a resulting user experience.

5.1.1. Content

Under this category, we grouped findings associated to participants' view on the process information and on the ability of the automation to deal with the process. We assume that the "content" addressed by robots is the programmed task(s) it executes. Also under content, we list the findings related to the process even if not automated, because we assume that the process is subject to being eventually "contained" in robots.

Participants reported that the process became more organized after robots. This happened because automation was preceded by a process analysis, reengineering and standardization. Standardization of activities was perceived as beneficial to the understanding of process.

However, processes are perceived as complex; because robots encapsulate this complexity, information about processes is perceived as lengthy and confusing when workers have to deal with it: "*Robots follow the tendency to make things easy, to bring*

a lot of information. This may be good, but it can also be bad because you may get lost among all the options you have" (S1). "*Lots of information all the time, all messed up, and we have to organize ourselves to understand each piece*" (S2).

Partially automated processes incur in workload being shared between robots and humans. Workers understand themselves as robot's redundancy, responsible for processing in case of failures. Robots alleviate the workload, but those workers allocated to the automated process still occasionally face a heavy workload, which is unpredictable because it may depend on external variables, such as the rainy season or calendar day, or contingency. "*When the server went down, the robot did not work. Then our work was a bit heavier, because then we had to deal with everything*" (S4).

Workers report they miss indicators of performance as an important piece of content. Present robot implementation lack visibility on its real performance compared to intended performance, as well as numbers that express false-positive and false-negative answers. "*We were surprised because what it was supposed to have done, it did not do. . . . When you lead a team, you keep an eye on the team performance. The robot is now part of the team, so I must keep track of its performance. And today this is not clear to us*" (S8). It is important to clarify that workers have a limited view of RPA KPIs such as execution rate, execution time and volume processed, which were defined and collected by the IT area responsible for the deployment of the robots.

5.1.2. Interaction

Under this category, we grouped those findings related to how participants describe the robot operability qualities, that is, how workers can control its operation.

From the workers' perspective, they understand themselves as robot's supervisors. They must keep track of the robot operation because in the end of the processing cycle, the robot informs the transactions it did not perform, so that workers can work on the remaining entries. "*The guy at the end is following the movement of the robot, so he knows that the robot has a routine. . . As it finished doing the shutdown, it starts an email or a log, for that user who is monitoring, [as if it said] 'look, I finished my activity'* " (S10).

5.1.3. Functionality

Under functionality, we categorized snippets that address the workers' perception on what the robot does or can do. Once G1 RPAs and humans work on the same process, the robots' functionality is well known to users. Our findings show robots not being allowed to fully perform their functionality.

Humans' work starts after the robot has finished processing and completes what robots could not manage, because humans have access to unstructured and complex information. The robot's speed to solve large amounts of transactions provides a positive experience, but because robots operate in a limited time window, sometimes it fails to complete their share of the work. Workers perceive that the robot is not as effective as it could be and that the robot operation could be managed by operational areas, who would tune its performance as needed. "*It still does not comply with 100%, there are many things that are its duty; however, it does not do it, so we end up having to do the work that was its task*" (S6).

Also, workers understand G1 as underused, for there are opportunities for more functionality and applications. "*I think there are improvements. . . I'm not even thinking about a future generation of robots, but I think this one can still be adjusted*" (S5).

5.1.4. Presentation

Under presentation, we categorized snippets that address the look and feel of the robot. Workers perceive presentation from their own perspective, as users, and from the clients' perspective.

From clients' perspective, workers understand that one of the robot's tasks is to communicate (via email) with clients. As cases processed by the robot issue immediate communication, workers assume that the robot's prompt communication provides a positive experience to clients. "*Our client contacts us because somehow he understands that our service caused him harm, he suffered a loss. The more agile I am in responding to him, the better it is*" (S7).

Workers express they were initially concerned with the clients' perception of robots. This concern is even greater with the new generation robots. Workers understand that clients' interaction with robots should be close to the interaction between humans. "*Will it run properly for clients? Will they notice? Or will it go undetected?*" (S1).

However, from the workers' own perspective, the robot is a black box, stealth by nature. Due to the complexity of business processes, workers miss visibility of the automated process rules. Although present workers know the process well, transparency in process execution can be beneficial to new workers and to process improvement. "*I wanted not only to push a button and get a result, but to have something that ensured that the robot actually performed the steps and did what needed to be done, so that it could guarantee the result it provided*" (S9).

Robots communicate the transactions it could not close by placing a lock in front of unprocessed entries. This icon is the key for humans understand that they can work on an entry because the robot has already checked it and was unable to solve it. An email is also sent with the information on the entries not processed. This regular communication is minimalist and effective for workers to observe that the robot is operating.

However, workers need indirect evidence to observe malfunctioning: the lack of communication indicates that the robot is inoperative. That is, in this case, supervision is done based on indirect evidence instead of clear status presentation. "*We only know that it is working because the amount [of entries]*

is decreasing [. . .] Then we check the count and the quantity is always the same, after half an hour, 40 minutes there is the same amount: Wait! There's something wrong" (S2). "*We do not have this contact, we do not see the robot acting, we do not see it working, we do not have control over the robots. I think that's what we need*" (S5).

Problems happen when workers spot errors in the robot operation, yet they cannot diagnose them, less likely act on correction. "*But it fails, we know that because the protocol is there. . . for some unknown reason the robot has not picked it, maybe because of the amount. . .* " (S6).

Robots thus fail in making their errors clear and visible.

5.2. Associated attributes

Under associated attributes, we compile those findings related to how workers perceive the system meeting their pragmatic and hedonic needs. Besides functionality and presentation, user experience with robots results in the perception of quality attributes, both present and desired. According to the Hassenzahl model, associated attributes can be of pragmatic and hedonic nature.

The Hedonic needs category expresses the perception of qualities associated with the satisfaction with the use and the to-be goals, while the Pragmatic needs category expresses the perception of usefulness and usability, and the to-do goals. In the Hassenzahl model, hedonic needs can be categorized into three groups: evocation, identification and stimulation. Pragmatic needs are categorized into usability and usefulness.

5.2.1. Evocation

Under evocation, we grouped those findings associated with the ability of the robot to stir participant's memories, and to represent values, events, relationships or thoughts that are important to the individual. Many different feelings were evoked by workers when talking about the robots, but undoubtedly, the most important value expressed by workers is the utility of released work time, for it provided an overall increase in life quality. The robot releases time previously used in repetitive tasks to more productive and pleasant usage. Also, it released workers from overtime, which was required when a high volume of work had to be processed to meet response deadlines: "*The robot releases my time and I can work on other things, dedicate myself to groups, make other interactions, look for new things*" (S1).

Workers were distressed and hopeless by the work before the robot was implemented. Back then, excessive workload came from different causes: from increased demand, resulting of system failures (anomalies process) or thunderstorm season (damage compensation process); from manual comparison of several data screens of non-integrated systems; from repetitive work that required focus and attention. In comparison to that scenario, the most common perception, expressed by all the participants, is that the presence of the robot evokes calmness, tranquility and comfort that is not translated into a complete relaxation, but instead as a positive feeling of engagement: "*at the same time I have to pay attention to the process and the information it is sending me, I can be a bit calm because I can count on the robot*" (S8).

Due to that, failure by the robot brings back the feelings of tedious, unpleasant work: "*We divide the work, a stage for each employee and we stay focused practically all day on the same thing, a very operational work. . . All in all, it is quite dull*" (S5).

The robot evokes the feeling of confidence in the process structure and on its work. Workers perceive the robots as trustworthy colleagues: "*I think our tendency is to trust their work. To go blind, that's what we did. Today, we do not look at what it's doing or worry about if it's doing right or wrong. [...] The more you know, the more you trust*" (S2).

In the beginning of their adoption process, robots evoked concern and surprise. Workers were concerned about the results of implementing the robot but were positively surprised by the robot being able to process over 70% of the entries, in the first days after its launching. "*It was a surprise because, although we expected it to process 70%, 80%, I did not expect that in the early days*" (S3).

Feelings of loneliness and remoteness were also reported. The work with the robot is perceived as distancing people because the person works with no human collaborator and because the robot is not physical. "*I do not miss the contact with the robot, I really miss human contact. I am the kind of person who likes human contact, so if I have been developing an activity for a long time without having an interaction with another person, I miss that*" (S9).

Robots evoke the expectation of unlimited possibilities of growth and new opportunities: "*It can evolve more ... We can achieve much more*" (S3).

5.2.2. Identification

According to Hassenzahl, "people also express their self through objects" (Hassenzahl et al., 2003). We grouped those findings related to how the work with robots contributes to users' identity or desired identity under this category.

We observed evidence of workers identifying themselves with the robots' maturity growth. Participation in a successful technological intervention brought a positive feeling of belonging and co-responsibility but also of superiority as we detail hereinafter.

Workers feel that they belong to the same organizational structure as the process and the robot. If the process gets more organized and defined, they feel professionally rewarded. "*We are part of the process; we end up being inside the process. If it is a messy process, if it is not going well, it automatically reflects on*

the professional. If you are in a process that people understand as well structured, that gives return or provides data that is relevant to the company; this is a gain for us as professionals for the company" (S5).

Owing to the impact of automation on jobs, workers feel responsible for the development of new opportunities. "*If I cannot prepare people for other things, if I do not provide new jobs, new opportunities, this can lead to people getting into a situation that we have caused, of unemployment, of having nothing to do*" (S1).

Working with robots has awakened feelings of maturity growth in the teams. Robots were observed in the first days from release. While the robot's maturity grew, also did the workers' trust in its work and in their own: "*It was something that we built and is solid, right?! We see that it is solid*" (S7).

Workers perceive they are still smarter than robots. The ability to infer situations and make quick decisions based on patterns makes human work faster than robots', which take longer to analyze all the applicable business rules. "*The robot processes the invoice item by item, that is, it takes a longer time than a person to analyse. Because the person already has the expertise, he/she takes a look, and he/she knows what to do*" (S4).

Due to the successful implementation of the robot, the consequent praise received from other companies and internal groups, and because of the active participation in the implementation process, workers feel intensely proud of their achievement. "*I managed to reverse the signal, in the sense that I previously had to beg for [RPA] to enter an area, now the area comes [to ask for RPA]. Then you see that it progressed a good way*" (S10). Even those workers who were displaced from repetitive work to higher level activities also feel proud of their career. "*Through the implementation of the robot, I was able to get here. It is a feeling of victory, of success*" (S3).

5.2.3. Stimulation

Again, according to Hassenzahl, people strive for personal development. When products can support this development, they have a stimulating effect. We categorized findings about the stimulation experience, both positive and negative, under stimulation.

We identified factors for stimulation in the change from tedious and repetitive work to opportunities of personal improvement. Also, stimulus come from knowledge of the process facilitating increased control and potential influence in the company. Those released from monotonous activities are stimulated by their deeper understanding of the company's business process and structures and their ability to act. "*Being able to act in these causes [of mistakes] is a bit complicated, and at the same time stimulating*" (S3).

They refer to this stimulus as invigorating: "*I feel invigorated by the opportunity to get out of something repetitive and do something new and think differently*" (S7).

Workers see that working with automation represents a professional advantage. RPA is a trend and this knowledge favors their employability. "*People who work in this area of robot development, AI, they will always be prominent because that is what the market was looking for*" (S2).

In the daily activities, however, work can be frustrating when the robot does not perform the tasks it was intended to: "*Sometimes it's frustrating… you expect it to do [the work] and it does not, it's kind of frustrating*" (S6).

5.2.4. Usability

The perception of usability is one of the two categories under the perception of pragmatic qualities. We classified reports of users' perceptions on classic attributes of usability under this category.

Workers perceive that the robot's execution is free from errors. "*So far, honestly, from the part of its work that we have analysed, I have not identified any flaws, no mistakes. About the work it performs, I have no doubt that it is a good job*" (S5).

The robot is fast; it is perceived as an efficient machine that will evolve to faster, more efficient: "*Its speed, the capability to execute […] and thinking that he might have an ever greater capacity, I think it will get more and more efficient, faster*" (S1).

Robots are perceived as effective and resource efficient. "*Robots are helping us, leveraging production with more effectiveness and a little better quality*" (S8). "*The technology works much faster and without errors, with fewer people - one managing the tool more than running [the process].*" Robot's processing was observed to be accurate; however, in some cases, accuracy is a source of issues. "*Many of, let's say, of the mishaps we had with the robot is the value review. Sometimes, on behalf of a cent the robot will leave [the entry] for the exception*" (S2).

Owing to robotized process quality, other processes are impacted. Workers perceive the overall process quality as improving after the deployment of the robot. "*After this process [PID] is finished, the customer can make a complaint. [...] So we also monitored the complaints process. And they have decreased in function of the quality that I applied to the other stages*" (S1).

5.2.5. Usefulness

Under this category, we classify snippets that address the perception of robot as a useful tool. Because robots add to the teamwork, they are seen as teammates. "*I understand that the robot is like a collaborator, it gave us more strength, more agility, so I think that looking from a global way it is adding, it is joining forces, assisting all the employees involved in the process*" (S5).

Usefulness is compromised by business changes. Robots are not expected to solve all the different cases; however, they are also expected to continuously expand their functionality and process coverage. Robots must evolve and also follow the

changes. "*Nowadays, we have a lot of exceptions, there's a lot that it tried to deal with and that it could not*" (S2).

Yet, robot improvement depends on the managerial decision to invest. "*We did it, developed it and we're just using it. In more than a year, we did not move. It will depend on whether someone wants to invest in it*" (S2).

5.3. Anticipated consequences

The model places attractiveness and behavior changes as categories of anticipated consequences, meaning that users, after their experience with the technology, may express likeliness of acceptance and anticipate their own future. In this study, the experience itself is not anticipated but cumulative, after 18 months of RPA adoption. Yet, participants have expressed their viewpoints about acceptance and their perception of the future, in the perspectives of technology evolution in the company and their personal destiny.

Participants foresee the broadening scope of robotization, with the robots working on non-digital or non-integrated sources of relevant data and the use of intelligent robots in more complex situations. These expectations are coherent with the newer generations of RPA. "*I believe that in the future we will be closer to this situation because, with the improvement of robots, they will do exactly what we want, help us to produce more, identify more mistakes, work more closely with us*" (S8).

In a more distant future, they believe that robots may evolve to resemble humans "*Because from what I see today of the robots being developed, it more and more tends to meet the personal needs of the human being and is becoming more and more like the human being*" (S9).

As expected, workers show concern about job loss in the future "*With robotization, you reduce labor. I believe that with fewer people, with the help of robots, you will end up decreasing your number of people*" (S6). However, they express their innocence of the unemployment process as they are expected to have pride in their accomplishment. "*My expectation is that the team, the people who use it, present it with satisfaction, saying 'look, this is what we did,' not with that other concern that you took a job, took space ...*" (S7).

Workers perceive digital technologies as agents for changes in people behavior. On the one hand, the change environment stimulates an attitude of creative restlessness, in which stagnation and apathy are not welcome. "*Even by the restlessness that I think we must have. When you've just seen one thing to improve you have to think about the next one, right?!*" (S7). On the other hand, they foresee an increase in interpersonal distance: "*I think people are losing some of their humanity, of this human coexistence*" (S1). Change in careers is also expected as workers realize that new-generation robots will promote people growth due to innovation.

5.4. Complementary categories

The analytical framework proposed in the AXE protocol contains three categories to group common statements in which participants reveal their position toward the product (unwanted and suggestions) or some important information that does not address the concept under evaluation or the activity directly linked to it (meta). Once situations perceived as undesirable could be associated with perceptions of features or qualities, they were classified under their respective categories and category unwanted was suppressed from this analysis.

5.4.1. Suggestions

Several suggestions for improvement were collected from participants' interviews, which are useful for future expansion. Workers perceive other time-consuming activities that could be replaced by a robot, freeing more time for more valuable activities. They also suggested that the management of robots' performance should be in the hands of the operational team, which could more efficiently manage the robot's schedule. "*If the scheduling was our task, because we know the amount every day, [we could change the schedule] and I think we would be adding even more value*" (S5). Also, participants expect more frequent revisions in the operational procedures. "*Because of these changes in procedures that happen all the time, I think we should periodically stop and check if something new has come up that we can include in the robot's activity, for example, to relieve some of the effort*" (S2).

5.4.2. Meta

Regarding the meta category in AXE framework, our research found several statements that did not address the usage of the robot or its intrinsic characteristics, but instead, reveal a lot about the implementation and deployment process.

The richness of these findings has motivated us to extend our understanding of the UX concept to encompass managerial factors, which show that UX is strongly affected by contextual factors and by the collectivity of workers that share the organizational situation. In order to guide our understanding, we considered Juntunen's integrative framework categories, which are presented in the following section.

6. Results of acceptance and adoption

Juntunen's integrative framework for acceptance and adoption, presented in Section C, has guided our understanding of the contextual factors that conditioned user experience [21].

She explains that an RPA decision on adoption is a top-down movement (the primary adoption) which is followed by

a series of actions toward internalization or second adoption. Acceptance, understood as workers' willingness to adhere to the adoption process, is influenced by several factors and beliefs. Her framework was useful to show the complex relationship between the contextual factors and their impact on user experience. We present the definition of each contextual factors and detail their manifestation in our study in the following sections.

6.1. Innovation attributes

The first category in Juntunen's model is the set of innovation attributes. Innovation attributes express constructs associated to the innovative characteristics of the product, service or idea being adopted. Innovation attributes can be closely related to the UX constructs of perceived features and associated attributes, especially those related to pragmatic qualities.

6.1.1. Relative advantage

Relative advantage reflects the superiority of the innovation over the previous idea it is replacing. There is a clear superiority of the robots over manual work, not only because of the released work time but also because of the increased productivity. Workers express that competitivity demands a quality leap that does not depend on hiring people. "*I had collaborators with potential for a larger delivery or for participation in some projects. I could not allocate them because they were servicing the high-demand process. So, to have them participating we had to wait for the low-demand season*" (S1). "*We already had the desire to gain speed with repetitive processes, to become more and more competitive. There is a cost issue. Sometimes, to do more, you will not achieve this by hiring more people*" (S7).

6.1.2. Complexity

Complexity refers to the perception of easiness to understand or use the technology.

The perception of technology being easy to use and to understand is strongly related to constructs of usability and user experience. Complexity is not perceived as associated to RPA but to the business process, and the RPA technology is perceived as a tool to reduce complexity, as processes are standardized before being automated. However, collaboration with the robots is affected by the lack of visibility of the robots' actions. Workers must develop strategies to cope with the lack of information on robots' work. "*As the time span for the robot to do this analysis may not have passed, we always pick [entries from] the day before, so we are sure that the robot was able to do the analysis of all requests*" (S8).

6.1.3. Trialability

Trialability is associated to how easily individuals can experiment with the innovation, thus, this category groups snippets that express how workers have experimented with the robots. Some of the participants had the opportunity to observe the robots from the first days of release. While the robot's maturity grew, also did the workers' trust in its work."*[Immediately after robot deployment] we would take everything it had processed the day before, and we checked if it had processed correctly, even to make the corrections at the beginning of the implantation. There's always something to be done, right? And then, from the moment we saw that its margin of error was very low, 1% or 2% of the amount that came in, we did not have the need to make this verification ... [...] So we spent a month checking if what it was doing was correct*" (S2). Trialability was an important factor for acceptance, because several adjustments had to be made due to the several exceptions to business rules. "*On paper it is one thing, but the moment you put it to work, you are faced with various situations. But I do not say it's luck, no, I believe it's trial and error. And then, you fix it and do it again until you adjust it the way you want it*". Modifiability is also perceived as an important factor, because workers observe that changes in business rules must be implemented quickly. "If you take too long to make [the adjustment] that you have identified, it becomes obsolete" (S7).

6.1.4. Observability

Under observability, we classified those snippets that express the workers' perception of the innovation being visible to others.

Robots are observable and demanded internally. Workers report that RPA is demanded from other business areas. "*I managed to reverse the signal, in the sense that I previously had to beg for [RPA] to enter an area, now the area comes [to ask for RPA]. Then you see that it progressed a good way*" (S10).

Outside the company, however, the technology is not visible by clients. Clients are unaware of changes and only notice its effects. Changes are then open to interpretation, as is the case when a client claimed the company did not analyze her case because the answer came too soon: "*They [the clients] said "the proof that you do not even check if there has been an incident in the grid: now you respond within one workday*" (S1).

6.1.5. Job-fit

Under this category, we grouped the findings associated to the perception that the innovation is compatible with the job it must perform. This is a relevant factor for acceptance in this case. Findings on associated attributes in the user experience framework showed that, in the early days of adoption, workers perceived robots' work as reliable and effective, freeing valuable time. "*It theoretically does not do the wrong thing. What it is intended to do, it does, and does it well. What is assigned for it to*

do, it can do and it does" (S6). "*When the robot was deployed, we had a very good time of having little work to do*" (S2). However, job-fit can degrade if the robot is not updated and improved. "*[The robot] solves the easy cases, at least for now*" (S4). "*The exceptions that we could not map are not processed by the robot, and this demand is passed on to us*" (S5).

6.1.6. Voluntariness

This category includes the findings related to the innovation adoption being perceived as voluntary. Workers perceive that RPA adoption is a top-down, irreversible movement. "*This is the first feeling we have when it comes to information that you will participate in a process to robotize your activity. Like a bitten apple*" (S3). The technology is accepted as the innovation is seen as an opportunity for evolution. "*It brings the feeling that I can improve, that I can move forward*" (S3). "*One has to try and reinvent oneself. Otherwise you, in fact, will be left behind*" (S10).

6.2. Organizational attributes

The organizational context influences the adoption process. All the attributes in Juntunen's framework were found to be relevant, but two additional organizational attributes were observed to be relevant to the case and were added to this analysis: perception of impact on clients and process improvement.

6.2.1. Compatibility as social responsibility

The compatibility attribute should reflect the innovation matching organizational needs and values. We found that workers praise the value of social responsibility in automation, which mirrors the goals in the Pact for Humanized Work Digitalization.

Workers understand that automation raises the level of requirements for hiring. "*Preparing for the digital world is a fairly complex social issue as to what is required in the labor market. If today we fail to supply the labor market with people for repetitive activities, let alone for technology*" (S1).

6.2.2. Organizational norms

The organizational norms attribute reflects organizational and leaders' attitude toward change. Competitivity is seen as one of such factors. Workers express that competitivity demands a quality leap that does not depend on hiring people, and the adoption of robots is a question of embracing competitivity. "*We already had the desire to gain speed with repetitive processes, to become more and more competitive. There is a cost issue. Sometimes, to do more, you will not achieve this by hiring more people. So I think it was a true evolution; this deployment was positive. This specific case was so successful that we presented it and as I said, other companies came to ask how it was done*" (S7).

6.2.3. Innovativeness

Innovativeness reflects the perception that the organization is early in adopting the innovation. We found organizational innovativeness to be a relevant factor as workers are proud of the company being an early adopter of RPA in Brazil.

6.2.4. Resource factors

Resource factors reflect the availability of money, skills and cooperation. Under this attribute, we placed snippets related to the company's investment in the innovation process.

Workers are aware of the company investment in RPA and of the requirement of return. Workers acknowledge that technological advances must be economically justified. "*[RPA] was a bet we all made. We focused on our results and the robots brought this to us. We bet on a machine process and it was not a roulette game because it was well thought out and we were sure we would make a profit on it*" (S7).

Workers understands that RPA provides no FTE reduction but a change in duties and avoiding increased costs with labor. "*Sometimes the investment is not just to add, but to maintain as well*" (S7). "*We often failed in the process due to the lack of workforce, due to the lack of people. Today because of robotization, we can execute all the steps*" (S5). "*If I am reducing man-hours, I am also reducing overtime. So it has a financial impact that is also expected by the company*" (S10).

6.2.5. Impact on clients

As to any company, the interface with customers is sensitive for EdP and dealt with extreme care. Besides image issues, failures in this interface may result in fines imposed by the regulatory agency. Participants expressed their concern with the impact of technology adoption on their clients. In the studied cases, the interface from robot to customer are emails that report customer's request status. It is important to notice that similarly to the workers' experience with RPA, the interface is not solely responsible for the customers' experience. The effect of automation on processes outputs also produces impact on customers' experience. In general, they expect clients to be positively affected by RPA. "*If it [the robot] finds an incident, it immediately reports it to the client [that it found an incident with a causal link]*" (S1) "*So we improve agility. It adds value to the company and to the consumer as well. If you have damaged equipment in your home [...] without knowing if you can fix it ... it is painful [...]*" (S5). "*... the company is not so susceptible to error, since a person can err more than a robot*" (S8).

Conversely, workers acknowledge that because consumers are not aware of robot presence or do not understand their operation, they may misunderstand agility as negligence. *"They [the clients] said "the proof that you do not even check if there has been an incident in the grid: now you respond within one workday""* (S1). Also, workers perceive that depending on socioeconomic factors, clients may experience difficulties in dealing with robots. As difficulties can come from clients not understanding the company's rules well or their own rights and duties in the contract, possibly such clients would benefit from human service. "*I see there's a chance we can use robotization to get closer to our customers. But you must choose well which technology will be used, and how we will use it. I think that [robots] should be used to get to know our customer well, but to serve the customer, I think maybe not" (S8). "Because we did not clearly disclose the information to our clients, that bad situation [the client experienced], he will keep it for the rest of his life*" (S6).

6.2.6. Process improvement

Process improvement was also perceived as an additional organizational factor, relevant in the studied case.

Workers attribute several improvements on business processes to RPA adoption. Improvement was clearly an outcome of the standardization needed for programming the robot activity but it is also reported to be a consequence of workers having free time to analyze the business process and its exceptions. As analytical workers, they can carry out their analysis and improvement beyond the context of one process and begin to look at relationships and mutual influence between processes. RPA is reported to promote synergy of workers toward process improvement. "*When we get a fault that is not the robot's, I check what the problem in the process was, act on the root cause and make the correction*" (S3). "*[Out of the critical period of operational work] we can work more analytically. We can give more attention [to the process], apply improvements, analyse the data the process generates, we can share this data with other areas of the company. The quality of the process has improved significantly*" (S5). "*So an improper handling of an anomaly will lead to a damage compensation complaint. So today I go deep into that complaint to see what happened. When I worked on the billing anomaly process I couldn't see that I was making a mistake, so I could correct it*" (S3). "*Because with the repetitive process being done by the robot, it created an opportunity for people to be closer and to discuss more about the process*" (S7).

6.3. Individual attributes

Individual attributes express personal characteristics influential to the process of change. Some of the attributes were not verified in our interviews: personality, age, gender and communication behavior.

6.3.1. Socioeconomic factors

Socioeconomic factors were expected to influence RPA adoption, as automation is often seen as a cause of workforce reduction. In the studied case, RPA was first seen as a workforce reduction policy, but participants reported that this idea has changed over time. "*When robotization started, that was my feeling here, it was taking a piece out of me, out of my activity. I thought that way. [...] Damn, I'm going to lose my job*" (S3). "*We know there is the fear of robotization threatening job positions and everything else, but I see the robot actually as a member of staff*" (S7).

Job vulnerability manifests itself in different ways, as pressure for productivity and dissatisfaction with the present national political and economic scenario: "*Because the tendency of all companies is to have fewer people doing more work. So you will always be more demanded for more activity. [. . .] It is not only in this company, but in the market, in general*" (S2). "*There's a lot of stuff involved, even the situation in the country. Now, maybe, we will not even retire. We're going to work for the rest of our lives*" (S2).

6.3.2. Innovativeness

Individual innovativeness is seen as a relevant factor in the studied case. Workers express that employability values are changing from knowledge background to innovativeness. "*I think that for you to deal with technology, for you to deal with machine development, you have to have a gift ... You must like it, you must study hard, but you must have great creativity. [. . .] That view, "Oh, you will only get a good job if you have a college degree" is changing. I think now it is: "if you show that you have something different*" (S2).

RPA is not the first automation technology as end-user programming was found to be practiced before robots' adoption. "*The staff ends up using Visual Basic language within Excel and do many things to automate our process*" (S7).

6.3.3. Expertise

Individual expertise was not seen as a factor for adoption or acceptance of RPA, but as a factor for keeping up with future advances in technology. Workers appreciate the fact that working in the company forces them to catch up with technology trends. "*If you're in the company, you end up adapting to technology and keeping up with growth*" (S4).

6.4. Managerial facilitation

Managerial facilitation describes the management approach to facilitate and to accelerate the adoption. Managers and technical leaders exert influence over workers. In the present case, according to other cases reported in the literature,

RPA adoption was decided at higher management levels and negotiated with the business areas. The snippets under this category describe the dissemination of the innovation idea throughout the company social structure.

6.4.1. Active participation

This category includes workers perception on their participation in RPA design.

The adoption process is clearly top-down, but workers see that RPA is a top-down movement that becomes accepted as people actively participate. RPA was first seen as a workforce reduction policy but as an inevitable progress in all companies, workers also expect to be under pressure for productivity. "*When robotization started, that was my feeling here, it was taking a piece out of me, out of my activity. I thought that way. [...] Damn, I'm going to lose my job*" (S3). "*Because the tendency of all companies is to have fewer people doing more work. So you will always be more demanded for more activity.[. . .] It is not only in this company, but in the market, in general*" (S2).

This view changed over time, as workers perceived that employability values are changing from knowledge background to innovativeness, or creative restlessness as they called their feeling. "*Even by the restlessness that I think we must have. When you've just seen one thing to improve you have to think about the next one!*" (S7).

Undoubtedly, the most influential managerial attitude in adoption was the considerate engagement of professionals, providing active participation of the workforce in RPA definition, programming and deployment. Workers participated by providing relevant information for development and changes and then verifying the robots' results until the technology was perceived as mature. "*The end user was the main source of information for constructing this robot. So I think that's a key point because people advocate what they have the opportunity to participate in*" (S7).

Workers see that their participation is still needed for robot improvement. "*We just have to have time. [. . .]to monitor [the process], to stop and say: 'no, we can do this differently' and then, get in touch with the people that develop the robot*" (S2).

6.4.2. HRM practices

Human Resource Management activities proved influential in removing adoption barriers and in promoting extrinsic motivation. Workers report that human resources spared by robots are being applied to purposeful tasks, which are more effective for the company and more rewarding to workers. They see strategic thinking, technology and innovation as advantageous skills in the new work market scenario, and clerical, repetitive work as less valued. They see themselves as responsible for developing technological professional competence. "*There was a time when the market was looking for people. Today, not anymore. The market is looking for technology*" (S2). "*If you remove the robots, we will have to go back to all those operational, monotonous activities that require a lot of work and do not add much value, even to the professional. We cannot improve professionally by performing these activities*" (S5). "*It is characteristic of the company to invest in people. [. . . The company] provided that many people had a college degree. And what happens? These people want to get out of the operational positions and go to analyst positions. [. . .] Consequently, the actual operational activity, if possible, must be robotized, so that individuals have the opportunity to do the analytical work that will bring benefits to the company and to them*" (S7).

6.4.3. Management of information

This attribute is related to the channels for sharing information about the innovation.

Workers praise the considerate strategy for introducing innovation, which was anchored in straightforward and clear communication.

"*I think people need to be clear, right? [. . .] What helped me a lot was a conversation I had with the manager, who explained to me how things would happen and showed me that if I participated in the development process, it would be good for my career, I would get more visibility. And this really happened." (S3) "I am very proud to work here, of using the methodology that was used, because of the concern that the company had to take people from our team to learn how to use technology, to have these people participating in this construction process and everything. I think that the success we have here [relates] to the architecture this project had, not only the technical architecture, but using cutting-edge technologies, building a structure, having people focused and resources in case it didn't work*" (S7).

The initial communication strategy was successful, but workers are currently engaged in setting new communication channels that carry the necessary information for maintenance, monitoring and dissemination. "*With the implementation of the robot, I saw the need to be in touch with the business areas, asking for feedback [on the robot]. How's it going? Is it satisfying your area? Does it help you in your daily life?*" (S10).

Since there are similar groups in different Federation states, communication is seen as essential, yet workers perceive that communication to be still too reactive. "*We do [make some change] here and do not communicate there or the guy does there [in another State unit] and does not communicate here. He is not isolated; he just does not interact*" (S10).

6.4.4. Influence strategies

This attribute is related to strategies to informally influence workers toward adoption. Besides active participation, seen

as the most effective organizational strategy toward adoption, organizational alignment is also considered a relevant factor. RPA introduction requires disclosure of information from different perspectives. "*We are talking from the architecture to the availability of software, the involvement of investments so that people at the operational level also engage, they buy the idea. So I think there is a concept of working, of disclosing and aligning the expectation with everyone, which must involve the needs, from the top management, to those who are there in their routine*" (S10).

7. Discussion

In this section, we summarize our findings and present our reflection on their meaning, as well as on how to improve workers' experience with RPAs.

7.1. Summary of findings

Our research question was stated as "After 18 months of the decision to adopt RPA in EdP, how do workers characterize their experience with RPA technology?". This research question was decomposed into three secondary questions, which are answered by our results in the previous sections. Although the experience phenomenon is too complex for a Manicheistic evaluation, we compile those factors that can be accountable for a clear positive or negative experience.

7.1.1. Factor for positive experience

Workers perceive RPA positively because:

1. RPA promoted process reengineering and standardization of activities resulting in more organization and efficiency.
2. They see RPA as a worker under their supervision, thus they perceive themselves to be in control.
3. RPA prompt feedback to clients sends a good message of agility.
4. RPA communicates the results of its correct operation using a minimalist and objective language.
5. The time released by the RPA operation is valuable.
6. Workers experienced more tranquility in work after the robot implementation
7. Workers are free from repetitive work and can dedicate their time to more engaging activities.
8. Robots are trustworthy.
9. Robots were surprisingly effective from the first days.
10. Work with robotization makes workers more employable.
11. Workers belong to a successful team after RPA deployment.
12. Work with robotization makes workers more valuable to the company.
13. Better understanding of processes makes workers more valuable to the company.
14. Workers feel responsible for the development of new careers.
15. Workers feel they are still smarter than robots.
16. Robots are fast and reliable.
17. Robots are effective and resource efficient.
18. Robots are useful.

7.1.2. Factors of negative experience

Workers perceive RPA negatively because:

1. There is no visibility of the process executed by RPA.
2. Cooperative work between process and humans makes humans the robot's redundancy, thus resulting in unpredictable workload.
3. Workers miss performance indicators that can help faster identification of failures and unsolved cases in which human action is required.
4. Robots' operation in limited time windows results in not reaching full performance. In this case, the team suggest that robots should be managed by the operational team.
5. There are many other conditions and processes that could be allocated to robots.
6. RPA does not make its status visible.
7. RPA does not help diagnose operation errors which can result in fines.
8. Working with robots lessen human contact and increase loneliness.
9. Workers feel frustrated when they have to return to manual activities, mainly due to failures.
10. Robot's accuracy prevents it from closing issues.
11. RPA utility is compromised by business changes.
12. Robot improvement is dependent on managerial decision to invest.

7.1.3. Workers' expectation toward technology evolution

After RPA implementation, workers' expectations regarding the future of technology and their own destiny can be summarized as:

1. Robotization will address more processes and will integrate other sources of data.
2. Robots will become more intelligent.
3. Robots will have to behave as humans.
4. There will be job losses.
5. Workers will become more creative and active.

7.2. Reflecions on the findings

In this section, we reflect on the research findings. We observed, as expected, that the experience is not formed only with the contact with the technology but is strongly influenced by the context in which this technology is inserted. Thus, the combination of Hassenzahl's framework, which focuses on the experience with technology, and Juntunen's framework, that focuses on the adoption and acceptance of technology, together, promoted a valuable tool to understand the multifaceted experience.

It is interesting to reflect on the findings by the lens of Juntunen's construct of beliefs. Her research concluded that the adoption process is influenced through the beliefs of individuals. We notice that her five influential beliefs also helps in framing user experience.

First, **perceived benefits** were clearly related to positive experience and behavioral change. The most influential factor, expressed by the totality of participants, has been the utility of released time both for professional and personal purposes. As the robots freed them from extremely boring and discouraging tasks, they were able to envisage a better future of more challenging and rewarding work situations. We believe that this perception may have been beneficial for the introduction of new, more sophisticated RPA technology that followed the deployment the G1 RPA (Vajgel et al., 2021).

On the other hand, **perceived social pressure** was influential, but in different terms as defined in the model. The negative risk of being against the social tendency was not manifested because workers perceive themselves as pioneers in RPA adoption in their environment. Thus, social pressure is an influential factor considering early adoption as an opportunity or positive risk.

Perceived effort is clearly seen as influential, although the perception of effort and its consequences varied between participants. Participants that were released from tedious activities believed that the effort put in the development of robots was worth the consequences, but those workers who were allocated to complete the work left by the robots perceive that their individual resources were not spared by the technology. The negative experience with the opacity of RPA execution can be removed by the design of dashboards that presents in real time the robot execution status, operation schedule and reasons for non-performance. This design is not straightforward in the specific RPA technology and this is a relevant improvement opportunity.

Regarding **perceived need and appropriateness**, workers believe RPA was a cost-effective solution for the processes in which the technology was inserted. The negative experiences stemmed from the fact that process owners do not see the possibility of full appropriation of technology, appropriation seen as the capacity to captain the evolution of technology and its conformation to needs. They they realize that without proper ownership, the cost of deployment may be wasted. because process improvement and resulting changes are inevitable. Without the power of apropriation, robots would quickly become obsolete.

Just as the perceived effort, **perceived capabilities** were also influential but in different ways. While some participants expressed their capabilities being challenged by the technology, those who were actively engaged in the change believed the technology helped demonstrate their capacities.

8. Conclusions

A recent literature review pointed to several challenges in RPA research. Purposefully, research on human factors was left out from the list of research challenges: "We perceive these human aspects of RPA to be similar to other technology adoption challenges, which could be addressed by the plethora of prior and ongoing IT adoption literature; hence, we have not focused on these in our formation of the research agenda" (Syed et al., 2020).

However, user research has already been proved to be a valuable strategy to guarantee that the design is suitable to human needs, expectations, habits and organizational norms. Our methodology resulted in insights that can help the company to deliver better automation to their workers in the subsequent RPA projects (Vajgel et al., 2021), and we understand that it can be useful for other companies that value human-centered design.

8.1. Methodology, limitations and future work

Regarding the methodology used, we believe that the UX framework and Juntunen's model were complementary in guiding our understanding of the factors that conditioned user experience. The UX methodology and its analysis model provided a detailed view on workers' relationship with the technology, while the acceptance and adoption model helped organize and explain the organizational and managerial factors that underlie that experience.

As an exploratory, qualitative study, our findings are limited due to the number of respondents and their roles. An interesting extension of this study would be the confrontation of workers' opinion on automation pains and gains with that of high-level managers, who were not interviewed in this study. Also, we were restricted to two processes in EdP Brazil scenarios. However, our research method can be replicated in other facilities and with other stakeholders to provide a reliable photograph of the human factors involved in the experience with RPA.

We intend to broaden our view on RPA UX by exploring other scenarios and RPA technologies. However, because the nature of UX is contextualized, we will invest on methodologies for efficient analysis of specific situations. Despite the stated limitations, we believe that our effort toward understanding workers pains and desires can be useful for future implementations of RPAs.

Data availability statement

The raw data supporting the conclusions of this article are not readily available because they may identify the participants who provided the information. Requests for access to the data should be directed to the corresponding author and are not guaranteed to be granted.

Ethics statement

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. The patients/participants provided their written informed consent to participate in this study.

Author contributions

LF: investigation, methodology, writing-original draft, and supervision. PC, SA-S, ST, MS, and REQ: investigation, writing-review, and editing. VD: project administration, investigation, writing-review, and editing. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by the Brazilian National Energy Power Agency (ANEEL)'s R&D program. REQ was supported by São Paulo Research Foundation (FAPESP) under grant 2019/21693-0.

Acknowledgments

The authors would like to thank the Brazilian National Energy Power Agency (ANEEL)'s R&D program; grant #2019/21693-0, São Paulo Research Foundation (FAPESP).

Conflict of interest

Authors ST, MS, and VD were employed by the company EdP Brasil.

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

Publisher's note

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

References

Alberth, M., and Mattern, M. (2017). Understanding robotic process automation (RPA). *J. Finan. Transform.* 46, 54–61. Available online at: https://www.capco.com/Capco-Institute/Journal-46-Automation/Understanding-robotic-process-automation

Al-Rahmi, W. M., Yahaya, N., Aldraiweesh, A. A., Alamri, M. M., Aljarboa, N. A., Alturki, U., et al. (2019). Integrating technology acceptance model with innovation diffusion theory: an empirical investigation on students' intention to use e-learning systems. *IEEE Access* 7, 26797–26809. doi: 10.1109/ACCESS.2019.2899368

Brasil, da Saúde, M., and de Saúde, C. N. (2016). "Resolução nº 510, de 7 de abril de 2016," in *Diário Oficial da União*, 44–46.

Chen, J. Y., and Barnes, M. J. (2014). Human-agent teaming for multirobot control: a review of human factors issues. *IEEE Trans. Hum. Mach. Syst.* 44, 13–29. doi: 10.1109/THMS.2013.2293535

Corbin, J., and Strauss, A. (2008). Strategies for qualitative data analysis. Basics of qualitative research. *Techn. Proc. Dev. Grounded Theory* 3, 9781452230153. doi: 10.4135/9781452230153

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Q.* 13, 319–340. doi: 10.2307/249008

Devarajan, Y. (2018). A study of robotic process automation use cases today for tomorrow's business. *Int. J. Comput. Techn.* 5, 12–18. Available online at: http://www.ijctjournal.org/Volume5/Issue6/IJCT-V5I6P3.pdf

Ernst and Young. (2015). *Robotic Process Automation White Paper*. Rio de Janeiro: Ernst and Young Assessoria Empresarial Ltd.

FDIs, I. (2008). *9241-210 (2009). Ergonomics of Human System Interaction-Part 210: Human-Centered Design for Interactive Systems (Formerly Known as 13407)*. Geneva: International Organization for Standardization (ISO).

Fontanillo Lopez, C. A., Li, G., and Zhang, D. (2020). Beyond technologies of electroencephalography-based brain-computer interfaces: a systematic review from commercial and ethical aspects. *Front. Neurosci.* 14, 611130. doi: 10.3389/fnins.2020.611130

Gegner, L., and Runonen, M. (2012). "Anticipated experience evaluation handbook," in *Proceedings of 8th International Design and Emotion Conference*, eds J. Brassett, P. Hekkert, G. Ludden, M. Malpass, and J. McDonnell (London).

Goodrich, M. A., and Schultz, A. C. (2008). *Human-Robot Interaction: A Survey*. Boston, MA: Now Publishers Inc.

Group, I. C. A. (2017). IEEE guide for terms and concepts in intelligent process automation.

Hassenzahl, M. (2003). T"he thing and i: understanding the relationship between user and product," in *Funology* (Dordrecht: Springer), 31–42.

Hassenzahl, M. (2013). "User experience and experience design," in *The Encyclopedia of Human-Computer Interaction 2* (Dordrecht).

Hassenzahl, M., Burmester, M., and Koller, F. (2003). "Attrakdiff: ein fragebogen zur messung wahrgenommener hedonischer und pragmatischer qualität," in *Mensch & Computer 2003* (Wiesbaden: Springer), 187–196.

Hassenzahl, M., and Tractinsky, N. (2006). User experience-a research agenda. *Behav. Inf. Technol.* 25, 91–97. doi: 10.1080/014492905003 30331

Houy, C., Hamberg, M., and Fettke, P. (2019). *Robotic process Automation in Public Administrations*. Bonn: Digitalisierung von Staat und Verwaltung.

ISO, I. (2011). IEC25010: 2011 systems and software engineering-systems and software quality requirements and evaluation (square)-system and software quality models. *Int. Organ. Stand.* 34:2910. Available online at: https://www.iso.org/standard/35733.html

Ivančić, L., Suša Vugec, D., and Bosilj Vukšić, V. (2019). "Robotic process automation: systematic literature review," in *International Conference on Business Process Management* (Springer), 280–295.

Juntunen, K. (2018). *Influence of Contextual Factors on the Adoption Process of Robotic Process Automation (RPA): Case study at Stora Enso Finance Delivery*. Uppsala: Uppsala University. p. 96. Available online at: https://www.diva-portal.org/smash/record.jsf?pid=diva2%3A1223866

Lacity, M., Willcocks, L., and Craig, A. (2016). "Robotizing global financial shared services at royal DSM," in *The Outsourcing Unit Working Research Paper Series* (London: Cham).

Leshob, A., Bourgouin, A., and Renard, L. (2018). "Towards a process analysis approach to adopt robotic process automation," in *2018 IEEE 15th International Conference on e-Business Engineering (ICEBE)* (Xi'an: IEEE), 46–53.

Lewicki, P., Tochowicz, J., and van Genuchten, J. (2019). Are robots taking our jobs? a roboplatform at a bank. *IEEE Softw.* 36, 101–104. doi: 10.1109/MS.2019.2897337

MarketsandMarkets (2017). "Robotic Process automation market by process operation - global forecast 2022," in *Robotic Process Automation Market* (Northbrook, IL).

Montero, J. C., Ramirez, A. J., and Enríquez, J. G. (2019). "Towards a method for automated testing in robotic process automation projects," in *2019 IEEE/ACM 14th International Workshop on Automation of Software Test (AST)* (Montreal, QC: IEEE), 4–47.

Nauwerck, G., and Cajander, A. (2019). "Automatic for the people: Implementing robotic process automation in social work," in *Proceedings of the 17th European Conference on Computer-Supported Cooperative Work-Demos and Posters. European Society for Socially Embedded Technologies (EUSSET)* (Salzburg).

Österholm, M. (2010). "Beliefs: a theoretically unnecessary construct?" in *Sixth Congress of the European Society for Research in Mathematics Education. January 28th-February 1st 2009* (Lyon: Institut National de Recherche Pédagogique), 154–163.

Prietch, S. S., and Filgueiras, L. V. L. (2015). "Technology acceptance evaluation by deaf students considering the inclusive education context," in *IFIP Conference on Human-Computer Interaction* (Cham: Springer), 20–37.

Ratia, M., Myllärniemi, J., and Helander, N. (2018). "Robotic process automation-creating value by digitalizing work in the private healthcare?" in *Proceedings of the 22nd International Academic Mindtrek Conference* (Tampere), 222–227.

Rogers, E. M., Singhal, A., and Quinlan, M. M. (2014). *Diffusion of Innovations*. New York, NY: Routledge.

Roto, V., Law, E., Vermeeran, A., and Hoonhout, J. (2011a). "UX white paper: Bringing clarity to the concept of UX," in *Dagstuhl Seminar Proceedings*, eds J. Hoonhout, E. L. -C. Law, V. Roto, and A. Vermeeran (Dagstuhl: Schloss Dagstuhl – Leibniz-Zentrum fur Informatik).

Roto, V., Law, E. L. -C., Vermeeran, A., and Hoonhout, J. (2011b). "10373 abstracts collection - demarcating user experience," in *Dagstuhl Seminar Proceedings (DagSemProc)* eds J. Hoonhout, E. L. -C. Law, V. Roto, and A. Vermeeran (Dagstuhl: Schloss Dagstuhl – Leibniz-Zentrum fur Informatik), 1–26. doi: 10.4230/DagSemProc.10373.1

Rutaganda, L., Bergstrom, R., Jayashekhar, A., Jayasinghe, D., Ahmed, J., et al. (2017). Avoiding pitfalls and unlocking real business value with RPA. *J. Finan. Transform.* 46, 104–115. Available online at: https://www.capco.com/Capco-Institute/Journal-46-Automation/Avoiding-pitfalls-and-unlocking-real-business-value-with-RPA

Schellen, E., Bossi, F., and Wykowska, A. (2021). Robot gaze behavior affects honesty in human-robot interaction. *Front. Artif. Intell.* 4, 663190. doi: 10.3389/frai.2021.663190

Stople, A., Steinsund, H., Iden, J., and Bygstad, B. (2017). Lightweight it and the it function: experiences from robotic process automation in a norwegian bank. *Bibsys Open J. Syst.* 25, 1–11. Available online at: https://www.researchgate.net/publication/321319770

Syed, R., Suriadi, S., Adams, M., Bandara, W., Leemans, S. J., Ouyang, C., et al. (2020). Robotic process automation: contemporary themes and challenges. *Comput. Ind.* 115, 103162. doi: 10.1016/j.compind.2019.103162

Vajgel, B., Corrêa, P. L. P., Tóssoli De Sousa, T., Encinas Quille, R. V., Bedoya, J. A. R., Almeida, G. M. D., et al. (2021). Development of intelligent robotic process automation: a utility case study in brazil. *IEEE Access* 9, 71222–71235. doi: 10.1109/ACCESS.2021.3075693

Venkatesh, V., Thong, J. Y., and Xu, X. (2012). Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology. *MIS Q.* 36, 157–178. doi: 10.2307/41410412

Vishnu, S., Agochiya, V., and Palkar, R. (2017). Data-centered dependencies and opportunities for robotics process automation in banking. *J. Finan. Transform.* 45, 68–76. Available online at: https://ideas.repec.org/a/ris/jofitr/1593.html

Willcocks, L. P., Lacity, M., and Craigm, A. (2015). *The IT Function and Robotic Process Automation*. London: The London School of Economics and Political Science.

frontiers | Frontiers in Computer Science

TYPE Original Research
PUBLISHED 16 March 2023
DOI 10.3389/fcomp.2023.1113903

OPEN ACCESS

EDITED BY
Simona Aracri,
National Research Council (CNR), Italy

REVIEWED BY
Galena Pisoni,
University of Nice Sophia Antipolis, France
Mohammad Faiz Iqbal Faiz,
Tezpur University, India

*CORRESPONDENCE
Philipp Brauner
✉ brauner@comm.rwth-aachen.de

SPECIALTY SECTION
This article was submitted to Human-Media Interaction, a section of the journal Frontiers in Computer Science

RECEIVED 01 December 2022
ACCEPTED 28 February 2023
PUBLISHED 16 March 2023

CITATION
Brauner P, Hick A, Philipsen R and Ziefle M (2023) What does the public think about artificial intelligence?—A criticality map to understand bias in the public perception of AI. *Front. Comput. Sci.* 5:1113903. doi: 10.3389/fcomp.2023.1113903

What does the public think about artificial intelligence?—A criticality map to understand bias in the public perception of AI

Philipp Brauner*, Alexander Hick, Ralf Philipsen and Martina Ziefle

Human-Computer Interaction Center, RWTH Aachen University, Aachen, Germany

Introduction: Artificial Intelligence (AI) has become ubiquitous in medicine, business, manufacturing and transportation, and is entering our personal lives. Public perceptions of AI are often shaped either by admiration for its benefits and possibilities, or by uncertainties, potential threats and fears about this opaque and perceived as mysterious technology. Understanding the public perception of AI, as well as its requirements and attributions, is essential for responsible research and innovation and enables aligning the development and governance of future AI systems with individual and societal needs.

Methods: To contribute to this understanding, we asked 122 participants in Germany how they perceived 38 statements about artificial intelligence in different contexts (personal, economic, industrial, social, cultural, health). We assessed their personal evaluation and the perceived likelihood of these aspects becoming reality.

Results: We visualized the responses in a criticality map that allows the identification of issues that require particular attention from research and policy-making. The results show that the perceived evaluation and the perceived expectations differ considerably between the domains. The aspect perceived as most critical is the fear of cybersecurity threats, which is seen as highly likely and least liked.

Discussion: The diversity of users influenced the evaluation: People with lower trust rated the impact of AI as more positive but less likely. Compared to people with higher trust, they consider certain features and consequences of AI to be more desirable, but they think the impact of AI will be smaller. We conclude that AI is still a "black box" for many. Neither the opportunities nor the risks can yet be adequately assessed, which can lead to biased and irrational control beliefs in the public perception of AI. The article concludes with guidelines for promoting AI literacy to facilitate informed decision-making.

KEYWORDS

artificial intelligence, affect heuristic, public perception, user diversity, mental models, technology acceptance, responsible research and innovation (RRI), collingridge dilemma

1. Introduction

Artifical Intelligence (AI), Deep Neural Networks (DNN) and Machine Learning (ML) are the buzzwords of the moment. Although the origins of AI and ML date back decades, they have received a tremendous boost in recent years due to increased computing power, more available digital data, improved algorithms and a substantial increase in funding (Lecun et al., 2015; Statista, 2022).

While we are still a long way from Artificial General Intelligence (AGI) ("strong AI")—referring to an AI that matches human intelligence, and can adapt as well as transfer learning to new tasks (Grace et al., 2018)—it is undeniable that even "weak AI" and ML that focus on narrow tasks already have a huge impact on individuals, organizations and

our societies (West, 2018). While the former, aims at recreating human-like intelligence and behavior, the latter is applied to solve specific and narrowly defined tasks, such as image recognition, medical diagnosis, weather forecasts, or automated driving (Flowers, 2019). Interestingly, recent advancements in AI and its resulting increase in media coverage, can be explained by the progress in the domain of weak AI, like faster and more reliable image recognition, translation, text comprehension through DNNs and their sub types (Vaishya et al., 2020; Statista, 2022), as well as image or text generation (Brown et al., 2020). Despite the tremendous progress in weak AI in the recent years, AI still has considerable difficulties in transferring its capabilities to other problems (Binz and Schulz, 2023). However, public perceptions of AI are often shaped by science fiction characters portrayed as having strong AI, such as Marvin from The Hitchhiker's Guide to the Galaxy, Star Trek's Commander Data, the Terminator, or HAL 9000 from Space Odyssey (Gunkel, 2012; Gibson, 2019; Hick and Ziefle, 2022). These depictions can influence the public discourse on AI and skew it into an either overly expectant or unwarranted pessimistic narrative (Cugurullo and Acheampong, 2023; Hirsch-Kreinsen, 2023).

Much research has been done on developing improved algorithms, generating data, labeling for supervised learning, and studying the economic impact of AI on organizations (Makridakis, 2017; Lin, 2023), the workforce (Acemoglu and Restrepo, 2017; Brynjolfsson and Mitchell, 2017), and society (Wolff et al., 2020; Floridi and Cowls, 2022; Jovanovic et al., 2022). However, despite an increased interest in the public perception of AI (Zuiderwijk et al., 2021), it is essential to regularly update these academic insights. Understanding the individual perspective plays a central part since the adoption and diffusion of new technologies such as AI and ML can be driven by greater acceptance or significantly delayed by perceived barriers (Young et al., 2021).

In this article, we present a study in which we measured novices' expectations and evaluations of AI. Participants assessed the likelihood that certain AI related developments will occur and whether their feelings about these developments are positive or negative. In this way, we identify areas where expectations and evaluations are aligned, as well as areas where there are greater differences and potential for conflict. Since areas of greater disparity can hinder social acceptance (Slovic, 1987; Kelly et al., 2023), they need to be publicly discussed. Based on accessible and transparent information about AI and a societal discourse about its risks and benefits, these discrepancies can either be reduced or regulatory guidelines for AI can be developed.

A result of this study is a spatial *criticality map* for AI-based technologies that 1) can guide developers and implementers of AI technology with respect to socially critical aspects, 2) can guide policy making regarding specific areas in need of regulation, 3) inform researchers about areas that could be addressed to increase social acceptance, and 4) identify relevant points for school and university curricula to inform future generations about AI.

The article is structured as follows: Section 2 defines our understanding of AI and reviews recent developments and current projections on AI. Section 3 presents our approach to measuring people's perceptions of AI and the sample of our study. Section 4 presents the results of the study and concludes with a criticality map of AI technology. Finally, Section 5 discusses the findings, the limitations of this work, and concludes with suggestions on how our findings can be used by others.

2. Related work

This section first presents some of the most commonly used definitions of AI and elaborates on related concepts. It then presents studies in the field of AI perception and identifies research gaps.

2.1. Overview on AI

Definitions of Artifical Intelligence (AI) are as diverse as research on AI. The term AI was coined during the "Dartmouth workshop on Artificial Intelligence" in 1955. During that year's summer, the proposed definition of Artifical Intelligence (AI) was that "every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it" (McCarthy et al., 2006). In the year 1955—almost 70 years ago—researchers were convinced that—within a 2 month period—these machines would understand language, use abstract concepts, and could improve themselves. It was an ambitious goal that was followed by even more ambitious research directions and working definitions for AI.

AI is a branch of computer science that deals with the creation of intelligent machines that can perform tasks that typically require human intelligence, such as visual perception, speech recognition, decision-making, and language translation (Russell and Norvig, 2009; Marcus and Davis, 2019). ML, conversely, is a subset of AI that focuses on the development of algorithms and statistical models that enable machines to improve their performance on a specific task over time by learning from data, without being explicitly programmed.

A central introductory textbook on AI by Russel and Norvig defines it as "the designing and building of intelligent agents that receive percepts from the environment and take actions that affect that environment" (Russell and Norvig, 2009). The Cambridge Dictionary takes a somewhat different angle by defining AI as "the study of how to produce computers that have some of the qualities of the human mind, such as the ability to understand language, recognize pictures, solve problems, and learn" or as "computer technology that allows something to be done in a way that is similar to the way a human would do it" (Cambridge Dictionary, 2022). This kind of AI approximates the human mind and is built into a computer which is then used to solve some form of complex problem. On the one hand, this approach serves us with a well-defined line of events: We have a problem, develop a solution and, hopefully, will be able to solve the initial problem. The machine's job, or more precisely, an AI's job would be to find a solution, that is, give an answer to our question. On the other hand, this approach is rather narrow in scope. As Pablo Picasso famously commented in an interview for the Paris Review in 1964 "[Computers[1]] are useless. They can only give you answers." Picasso wanted to convey that a computer, or an AI for that matter, can only present outputs

1 Picasso was referring to mechanical calculation machines, nowadays called computers.

specific to an input, i.e., a specific answer to a specific question. However, it is currently beyond the capabilities of any AI algorithm to transfer its "knowledge" to any previously unseen problem and excel at solving it (Binz and Schulz, 2023). This is why there are many algorithms and many AI models, one for each particular problem. Going back to definitions—at least today—there is no single universal definition that captures the essence of AI.

Current AI research focuses on automating cognitive work that is often repetitive or tiring (Fosso Wamba et al., 2021). Its aim is to provide technological solutions to an otherwise inefficient or less efficient way of working. However, there are many other areas of (potential) AI applications that are merely an extension of what the human mind can do, such as creativity. In a recent example, a AI-based art generator won a prestigious art competition in the USA. In this case, the piece of art was entitled *The death of art* and received a mixed reception on Twitter, with some people fearing for their jobs, which may soon be replaced by a machine (Jumalon, 2022).

Many research articles focus on workers' perceptions of machine labor and its potential to replace some aspect of their work (Harari, 2017). In most cases, the machine is not a replacement, but rather an addition to the workforce (Topol, 2019). However, fear of replacement still exists among people working in jobs that are particularly easy to automate, such as assembly line work, customer service or administrative tasks (Smith and Anderson, 2014). A recent study found that workers' level of fear of being replaced did not significantly affect their level of preparation for this potential replacement, such as acquiring new skills. Furthermore, appreciation of the new technology and perceived opportunity positively influenced workers' attitudes toward automation (Rodriguez-Bustelo et al., 2020). This is just one example of the importance of perception of e.g., a new technology, and consequently its understanding, to accurately judge its implications.

Some form of AI is now used in almost all areas of technology, and it will continue to spread throughout society (Grace et al., 2018; Almars et al., 2022). Current application areas include voice assistants, automatic speech recognition, translation, and generation that can exceed the human performance (Corea, 2019), automated driving and flying (Klos et al., 2020; Kulida and Lebedev, 2020), and medical technologies (areas where AI could touch our personal lives) (Klos et al., 2020; Jovanovic et al., 2022), as well as production control (Brauner et al., 2022), robotics and human-robot interaction (Onnasch and Roesler, 2020; Robb et al., 2020), human resource management, and prescriptive machine maintenance (areas where AI could touch our professional lives).

We suspect that the perception of the benefits and potential risks of AI is influenced by the application domain and thus that the evaluation of AI cannot be separated from its context. For example, AI-based image recognition is used to evaluate medical images for cancer diagnosis (Litjens et al., 2017) or to provide autonomously driving cars with a model of their surroundings (Rao and Frtunikj, 2018). Therefore, people's perception of AI and its implications will depend less on the underlying algorithms and more on contextual factors.

2.2. Studies on human perception of AI

As outlined in the section above, perceptions of AI can be influenced not only by the diversity of end users (Taherdoost, 2018; Sindermann et al., 2021), but also by contextual influences. As an example from the context of automated driving, Awad et al. used an instance of Foot's *Trolly dilemma* (Foot, 1967) to study how people would prefer a AI-controlled car to react in the event of an unavoidable crash (Awad et al., 2018). In a series of decision tasks, participants had to decide if the car should rather kill a varying number of involved pedestrians or its car passengers. The results show that, for example, sparing people is preferred to sparing animals, sparing more people is preferred over sparing fewer people and, to a lesser extend, pedestrians are preferred to passengers. The article concludes that consideration of people's perceptions and preferences, combined with ethical principles, should guide the behavior of these autonomous machines.

In a different study (Araujo et al., 2020) examined the perceived usefulness of AI in three contexts (media, health, and law). As opposed to the automated driving example, their findings suggest that people are generally concerned about the risks of AI and question its fairness and usefulness for society. This means that in order to achieve appropriate and widespread adoption of AI technology, end-user perceptions and risk assessments should be taken into account at both the individual and societal levels.

In line with this claim, another study has investigated whether people assign different levels of trust to human, robotic or AI-based agents (Oksanen et al., 2020). In this study, the researchers investigated the extent to which participants would trust either an AI-based agent or a robot with their fictitious money during a so-called trust game, and whether the name of the AI-based agent or robot would have an influence on this amount of money. The results showed that the most trusted agent was a robot with a non-human name, and the least trusted i.e., the agents was given the least amount of money, was an unspecified control (meaning that it was not indicated if it was human or not) named Michael. The researchers concluded that people would trust a sophisticated technology more in a context where this technology had to be reliable in terms of cognitive performance and fairness. They also concluded that, from the Big Five personality model (McCrae and Costa, 1987), the dimension Openness was positively, and Conscientiousness negatively related to the attributed trust. The study provided support for the theory that higher levels of education, previous exposure to robots, or higher levels of self-efficacy in interacting with robots may influence levels of trust in these technologies.

In addition to this angle, the domain of implementation of AI, i.e., the role it takes on in a given context, was explored (Philipsen et al., 2022). Here, the researchers investigated what the roles of an AI are and how an AI has to be designed in order to fulfill the expected roles. On the one hand, the results show that people do not want to have a personal relationship with an AI, e.g., an AI as a friend or partner. On the other hand, the diversity of the users influenced the evaluation of the AI. That is, the higher the trust in an AI's handling of data, the more likely personal roles of AI were seen as an option. Preference for subordinate roles, such as an AI as a servant, was associated with general acceptance of technology

and a belief in a dangerous world. Thus, subordinate roles were preferred when participants believed that the world we live in is more dangerous than it is not. However, the attribution of roles was independent of the intention to use AI. Semantic perceptions of AI also differed only slightly from perceptions of human intelligence, e.g., in terms of morality and control. This supports our claim that initial perceptions of e.g., AI can influence subsequent evaluations and both, potentially and ultimately, AI adoption.

With AI becoming an integral part of lives as personal assistants (Alexa, Siri, …) (Burbach et al., 2019), large language models (ChatGPT, LaMDA, …), smart shopping lists, and the smart home (Rashidi and Mihailidis, 2013), end-user perception and evaluation fo these technologies becomes increasingly important (Wilkowska et al., 2018; Kelly et al., 2023). This is also evident in professional contexts, where AI is used—for example—in medical diagnosis (Kulkarni et al., 2020), health care (Oden and Witt, 2020; Jovanovic et al., 2022), aviation (Klos et al., 2020; Kulida and Lebedev, 2020), and production control (Brauner et al., 2022). The continued development of increasingly sophisticated AI can lead to profound changes for individuals, organizations and society as a whole (Bughin et al., 2018; Liu et al., 2021; Strich et al., 2021).

However, the assessment of the societal impact of a technology in general, and the assessment of AI in particular, is a typical case of the Collingridge (1982): These are developments that are either difficult to predict if they do not exist, or difficult to manage and regulate if they are already ubiquitous. On the one hand, if the technology is sufficiently developed and available, it can be well evaluated, but by then it is often too late to regulate the development. On the other hand, if the technology is new and not yet pervasive in our lives, it is difficult to assess its perception and potential impact, but it is easier to manage its development and use. Responsible research and innovation requires us to constantly update our understanding of the societal evaluations and implications as technologies develop (Burget et al., 2017; Owen and Pansera, 2019). Here, we aim to update our understanding of the social acceptability of AI and to identify any need for action (Owen et al., 2012).

3. Method

Above, we briefly introduced the term AI, showed that AI currently involves numerous areas of our personal and professional lives, and outlined studies on the perception of AI. The present study is concerned with laypersons' perceptions, their assessment of an AI development and its expected likelihood of actually happening. Thus, our approach is similar to the Delphi method, where (expert) participants are asked to make projections about future developments (Dalkey and Helmer, 1963), by aggregating impartial reflections of current perceptions into insights about technology adoption and technology foresight.

To assess perceptions of AI, we used a two-stage research model. In the first stage of our research, topics were identified in an expert workshop to get an accurate list. Then, these topics were rated by a convenient sample in the manner described above. This approach for studying laypeople's perception of AI will be further discussed later in the article.

3.1. Identification of the topics

To develop the list of topics we conducted a three-stage expert workshop with four experts in the field of technology development and technology forecasting. In the first stage, we brainstormed possible topics. In the second stage, similar topics were grouped and then the most relevant topics were selected, resulting 38 topics. In the third and final stage, the labels of the 38 defined topics were reworded so that they could be easily understood by the participants in the survey which followed.

3.2. Survey

We designed an online survey to assess non-experts' perceptions of AI. It consisted of two main parts: First, we asked about the participant's demographics and additional explanatory factors (see below). Second, we asked participants about numerous aspects and whether they thought the given development was *likely* to occur (i.e., Will this development happen?), and, as a measure of acceptability (Kelly et al., 2023), how they personally *evaluated* this development (i.e., Do you think this development is good?). Overall, we asked about the expectation (likelihood) and evaluation (valence) of 38 different aspects, ranging from the influence on the personal and professional life, to the perceived impact of AI on the economy, healthcare, and culture, as well as wider societal implications. The questionnaire was administered in German and the items were subsequently translated into English for this article. Figure 1 illustrates the research approach and the structure of the survey. Table 2 lists all statements from the AI scenarios.

3.3. Demographics and explanatory user factors

In order to investigate possible influences of user factors (demographics, attitudes) on the expectation and evaluation of the scenarios, the survey started with a block asking for demographic information and attitudes of the participants. Specifically, we asked participants about their age in years, their gender and their highest level of education. We then asked about the following explanatory user factors that influenced the perception and evaluation of technology in previous studies. We used 6-point Likert-scales to capture the explanatory user factors (ranging from 1 to 6). Internal reliability was tested using Cronbach's alpha (Cronbach, 1951).

Affinity for Technology Interaction refers to a person's "tendency to actively engage in intensive technology interaction" (Franke et al., 2019) and is associated with a positive basic attitude toward various technologies and presumably also toward AI. We used five items with the highest item-total-correlation. The scale achieved excellent internal reliability ($\alpha = 0.804$, $n = 122$, 5 items).

Trust is an important prerequisite for human coexistence and cooperation (Mc Knight et al., 2002; Hoff and Bashir, 2015). Mayer et al. (1995) defined trust as "the willingness of a party to be vulnerable to another party." As technology is perceived as social actor (Reeves and Nass, 1996), trust is also relevant to the acceptance and use of digital products and services. We used three

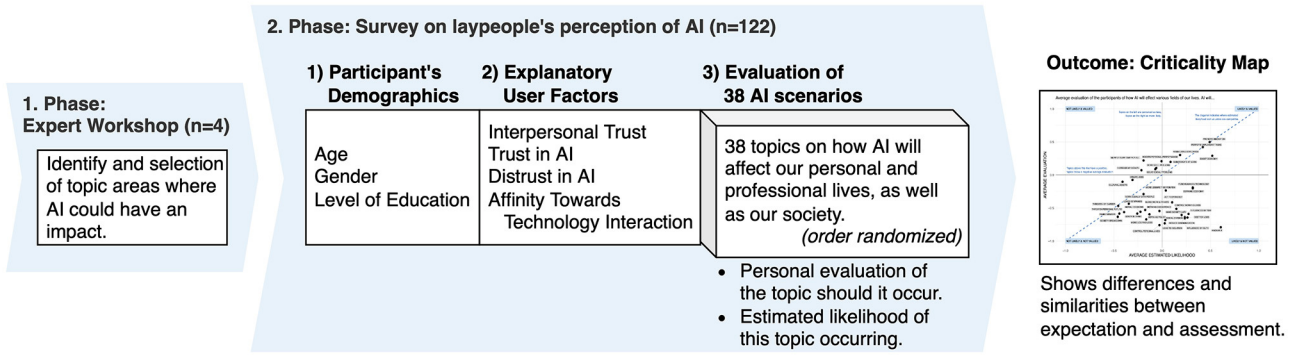


FIGURE 1
Multi-stage research design of this study with expert workshop and subsequent survey study. The questionnaire captures demographics, exploratory user factors and the evaluation of the 38 AI-related scenarios.

scales to measure trust: First, we measured *interpersonal trust* using the psychometrically well validated KUSIV3 short scale with three items (excellent internal reliability, $\alpha = 0.829$) (Beierlein, 2014). The scale measures the respondent's trust in other people. Secondly and thirdly, we developed two short scales with three items each to specifically model *trust in AI* and *distrust in AI*. Both scales achieved an acceptable internal reliability of $\alpha = 0.629$ (trust in AI) and $\alpha = 0.634$ (distrust in AI).

3.4. Perception of artificial intelligence

We asked about various topics in which AI already plays or could play a role in the future. The broader domains ranged from implications for the individual, over economical and societal changes, to questions of governance. Some of the topics were more straightforward and others rather far-flung.

For each of the 38 topics, we asked the participants whether this development is likely or not (*likelihood*) and if they evaluate this development as positive or negative (*evaluation*). Table 3 presents these topics that ranged from "AI will promote innovation," over "AI will create significant cultural assets," to "AI will lead to the downfall of society."

The questionnaire displayed the items in three columns: The item text on the left and two Likert scales to query the participants' expected likelihood and evaluation should the development come true on the right. The order of the items was randomized across the participants to compensate for question order biases. We used 4-point Likert scales to measure the expected likelihood of occurrence and evaluation of the given statements.

3.5. Survey distribution and data analysis

The link to the survey was distributed *via* email, messaging services, and social-networks. We checked that none of the user factors examined were correlated with not completing the survey and found no systematic bias. We therefore consider the dataset of 122 samples in the following.

We examined the dataset using the social sciences portfolio of methods (Dienes, 2008). To assess the association between the variables, we analyzed the data using non-parametric (Spearman's ρ) and parametric correlations (Pearson's r), setting the significance level at 5% ($\alpha = 0.05$). We used Cronbach's α to test the internal consistency of the explanatory user factors and, where permitted, calculated the corresponding scales. As there is no canonical order for the statements on the AI developments, we did not recode the values. We calculated mean scores (M) and the standard deviation (SD) for likelihood and evaluation for both the 38 developments (individually for each topic across all participants; vertical in the dataset) and for each participant (individually for each participant across all topics; horizontal in the dataset). The former gives the sample's average assessment of each topic, while the latter is an individual measure of how likely and how positive the participants consider the questioned developments are in general.

3.6. Description of the sample

In total, 122 people participated in the survey. Forty one identified themselves as men, 81 as women, and no one stated "diverse" or refused to answer. The age ranged from 18 to 69 years ($M = 33.9$, $SD = 12.8$). In the sample, age was neither associated with Affinity Toward Technology Interaction, nor with any of the three trust measures ($p > 0.05$). Gender was associated to Affinity Toward Technology Interaction ($r = -0.381$, $p < 0.001$), with men, on average, reporting higher attitudes toward interacting with technology. Interpersonal Trust is associated to higher Trust in AI ($r = 0.214$, $p = 0.018$), but not to higher *distrust* in AI ($p = 0.379$). Not surprisingly, there is a negative relationship between trust and distrust in AI ($r = -0.386$, $p < 0.01$. People who have more trust in AI report less distrust and vice versa. Finally, Affinity Toward Technology Interaction is related to both trust in AI ($r = 0.288$, $p = 0.001$ and (negatively) to distrust in AI ($r = -0.280$, $p = 0.002$). Table 1 shows the correlations between the (explanatory) user factors in the sample.

TABLE 1 Descriptive statistics and correlations of the (explanatory) user factors in the sample of 122 participants.

Variable	*M* (*SD*)	1	2	3	4	5
1. Age in years	33.88 (12.81)	–				
2. Gender	41 male, 81 female	−0.013	–			
3. Interpersonal trust	4.13 (0.96)	0.10	−0.06	–		
4. Affinity toward technology interaction	3.74 (1.16)	−0.15	−0.38	0.02	–	
5. Trust in AI	3.34 (0.87)	.06	−0.06	0.21	0.29	–
6. Distrust in AI	4.01 (1.00)	−0.01	0.09	−0.08	−0.28	−0.39

Note that gender is dummy-coded (0 = male, 1 = female).

4. Results

First, we analyse how participants evaluate the different statements on AI and map these statements spatially. Figure 2 shows a scatter plot of the participants' average estimated probability of occurrence and their average rating for each of the 38 topics in the survey. Each individual point in the figure represents the evaluation of one topic. The position of the points on the horizontal axis represents the estimated likelihood of occurrence, with topics rated as more likely to occur further to the right of the figure. The position on the vertical axis shows the rating of the statement, with topics rated as more positive appearing higher on the graph. Table 2 shows the individual statements and their ratings.

The resulting graph can be interpreted as a *criticality map* and read as follows: In the upper left corner are those aspects that were rated as positive but unlikely. The upper right corner shows statements that were rated as both positive and likely. The lower right corner contains statements that were rated as negative but likely. Finally, the lower right corner contains statements that were perceived as both negative and unlikely. Second, dots on or near the diagonal represent aspects where the perceived occurrence is consistent with the personal rating of the aspect: These aspects are either perceived as likely and positive (e.g., "promote innovation" or "do unpleasant activities" or as unlikely and negative (e.g., "occupy leading positions in working life" or "threaten my professional future"). On the other hand, for points off the diagonal, expectations and evaluations diverge. The future is either seen as probable and negative (e.g., "be hackable" or "be influenced by a few"), or as unlikely and positive (e.g., "create cultural assets" or "lead to more leisure time for everyone").

Accordingly, three sets of points deserve particular attention. Firstly, the points in the bottom half of the graph, as these are seen as negative by the participants. This is where future research and development should take people's concerns into account. Secondly, the points in the upper left quadrant of the graph, as these are considered positive but unlikely. These points provide insight into where participants perceive research and implementation of AI to fall short of what they want. Finally, all items where there is a large discrepancy between the likelihood of occurrence and the assessment (off the diagonal), as these items are likely to lead to greater uncertainty in the population.

As the figure shows, for some of the statements the estimated likelihood of occurrence is in line with the participants' personal assessment, while for others there is a strong divergence. The statements with the highest agreement were that AI will support the performance of unpleasant activities (positive expectation and evaluation), that it will promote innovation (also positive expectation and evaluation), that it will threaten the professional future of participants (both low evaluation and low expectation), and that AI will occupy leading positions in working life (again, both low evaluation and low expectation). In contrast, the statements with the largest difference share the pattern that they are expected to become reality and are viewed negatively by the participants. The statements were that the development and use of AI will be influenced by a few, that the use of AI will lead to less communication, that AI will be influenced by an elite, that it will destroy more jobs than it creates, and finally that it will be hackable.

4.1. Are the estimated likelihood of occurrence and the evaluation correlated?

Next, we analyse whether the expected likelihood and perceived valence ratings are correlated. To do this, we calculated Pearson's correlation coefficient between the average ratings of the 38 AI-related topics. The test showed a weak association of ($r = 0.215$), but this is not significant ($p = 0.196 > .05$). This means that expectations of potential developments are not related to people's evaluations of them. Thus, our sample does not provide evidence that the perceived likelihood and valence of AI's impact on society, personal and professional life are related.

4.2. Does user diversity influence the technology foresight?

Finally, we examined whether the explanatory user factors influenced the evaluation and estimated likelihood of the different AI topics. To do this, we calculated an average score for the two target dimensions for each participant. A correlation analysis shows that both the mean likelihood and the mean evaluations of the topics are influenced by user diversity and the explanatory user factors. Table 3 shows the results of the analyses. Across the participants the mean valence is weakly and negatively related to trust ($r = -0.253$, $p = 0.005$) and positively related to distrust in AI ($r = 0.221$, $p = 0.014$). Thus, participants with higher distrust in AI rated the potential scenarios as slightly more favorable, while higher trust is associated with slightly lower evaluations.

The mean estimated likelihood of occurrence is related with distrust in AI ($r = -0.336$, $p < 0.001$), Affinity Toward Technology Interaction ($r = 0.310$, $p < 0.001$), trust in AI ($r = 0.203$, $p = 0.025$), as well as to interpersonal trust ($r = 0.183$, $p = 0.043$). Higher distrust in AI is associated with a lower estimated likelihood, while all other variables are associated with a higher

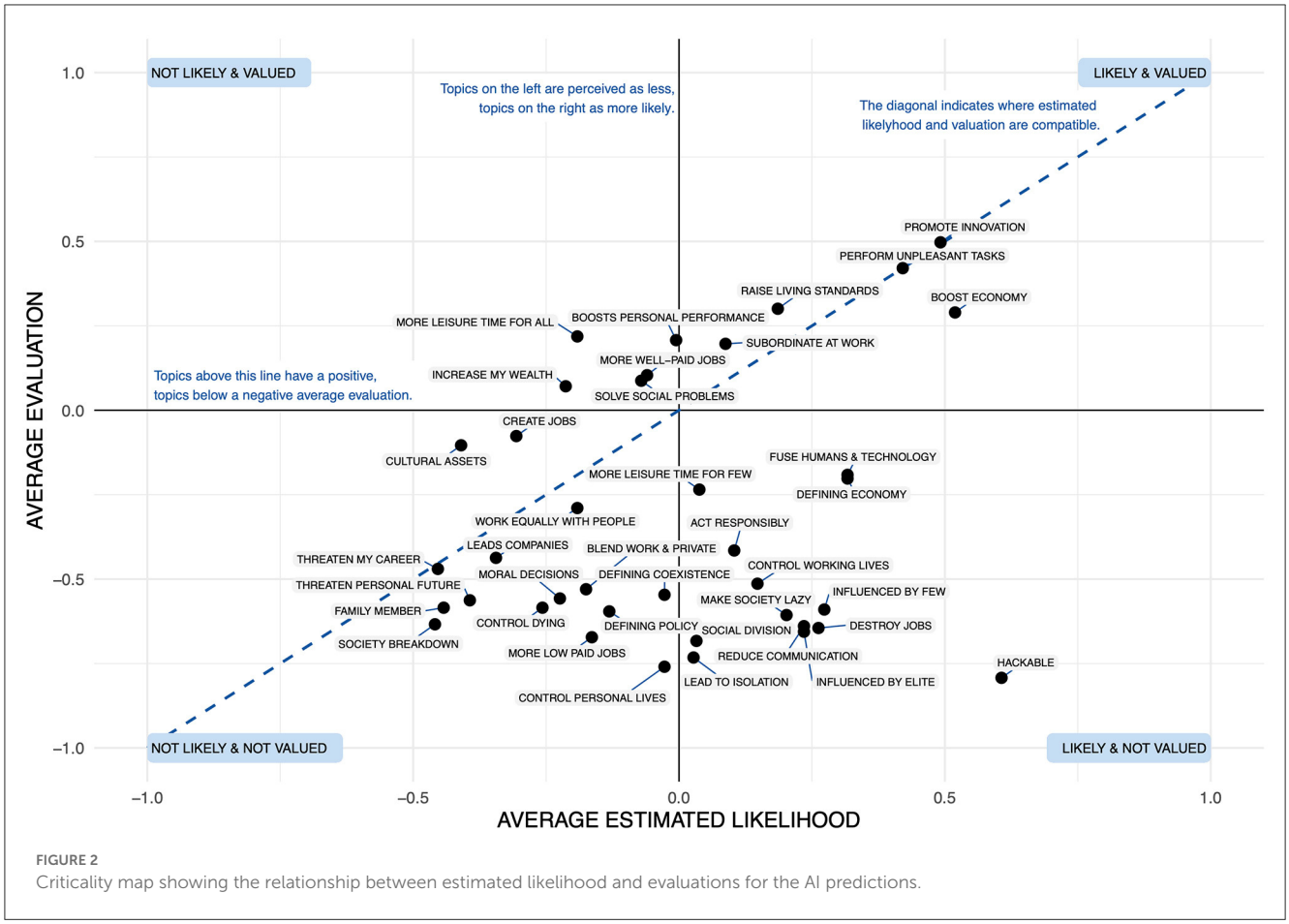


FIGURE 2
Criticality map showing the relationship between estimated likelihood and evaluations for the AI predictions.

estimated likelihood. It appears that distrust in AI is associated with a lower estimated likelihood.

5. Discussion

This article presents the results of a survey of people's expectations and evaluations of various statements about the impact AI might have on their lives and society. Overall, participants in our study associated AI with both positive and negative evaluations, and also considered certain developments to be more or less likely. Thus, AI and its implications are not perceived as either black or white, but participants had a nuanced view of how AI will affect their lives. From the perspective of social acceptance, issues of divergence between the two dimensions of expectation and evaluation deserve particular attention.

We analyzed the participants' subjective assessments of the developments. While this gives an insight into their beliefs and mental models, some of the assessments are likely to be challenged by other research. A critical point here is certainly the assessment of how AI will affect the labor market and individual employment opportunities. Our study participants are not very concerned about their professional future or the labor market as a whole. While a significant shift away from jobs with defined inputs and outputs (tasks perfectly suited for automation by AI) is predicted, which could lead to either lower employment or lower wages (Acemoglu and Restrepo, 2017; Brynjolfsson and Mitchell, 2017), participants in our sample do not feel personally affected by this development. They see a clearly positive effect on the overall economic performance and that in the context of AI it is likely that few new jobs will be created (and that jobs will rather be cut), but they do not see their individual future prospects as being at risk. This may be due to their qualifications or to an overestimation of their own market value in times of AI. Unfortunately, our research approach does not allow us to answer this question. However, comparing personal expectations, individual skills and future employment opportunities in the age of AI is an exciting research prospect.

Rather than examining the influence of individual differences, our study design focused on mapping expectations toward AI. However, this more explanatory analysis still revealed insights that deserve attention in research and policy making. Our results suggest that people with a lower general disposition to trust AI will, on average, evaluate the different statements more positively than people with a lower disposition to distrust. Similarly, a higher disposition to trust AI is associated with a lower average valence. When it comes to expectations for the future, the picture is reversed. A high disposition to trust is associated with a higher probability that the statements will come true, whereas a high disposition to distrust is associated with a lower expected probability. As a result, people with less trust rate the impact of AI as more positive but less likely. Thus, for this group, certain

TABLE 2 The participants' estimated likelihood (Likelihood) of occurrence and subjective assessment (Evaluation) of the various consequences AI could have on our lives and work[a].

AI will...	Likelihood		Evaluation	
	Mean	SD	Mean	SD
Do unpleasant activities	42.1%	56.8%	42.1%	64.7%
Promote innovation	49.2%	51.5%	49.7%	60.2%
Threaten my professional future	−45.4%	60.6%	−47.0%	66.9%
Occupy leading positions in working life	−34.4%	66.4%	−43.7%	65.3%
On equal footing at the workplace	−19.1%	58.9%	−29.0%	60.4%
Be subordinate in working life	8.7%	59.5%	19.7%	66.9%
Increase the standard of living	18.6%	59.1%	30.1%	64.6%
Become a family member	−44.3%	63.8%	−58.5%	59.6%
Solve complex social problems	−7.1%	62.1%	8.7%	66.5%
Lead to more well-paid jobs	−6.0%	67.4%	10.4%	69.5%
Threaten my private life	−39.3%	64.7%	−56.3%	60.5%
Lead to a downfall of society	−45.9%	59.6%	−63.4%	55.6%
Increase my personal performance	−0.5%	66.0%	20.8%	66.6%
Increase economic performance	51.9%	40.4%	29.0%	59.8%
Create more jobs	−30.6%	63.2%	−7.7%	75.9%
Lead to more leisure time for a few	3.8%	60.0%	−23.5%	62.8%
Increase my wealth	−21.3%	60.6%	7.1%	68.9%
Create cultural assets	−41.0%	55.7%	−10.4%	66.8%
Control our dying	−25.7%	64.8%	−58.5%	59.0%
Make moral decisions	−22.4%	70.9%	−55.7%	56.6%
Blend work and leisure time	−17.5%	59.4%	−53.0%	52.6%
Lead to more leisure time for everyone	−19.1%	55.7%	21.9%	66.8%
Defining political decisions	−13.1%	66.9%	−59.6%	52.5%
Lead to more low-paid jobs	−16.4%	75.0%	−67.2%	50.9%
Fuse humans and technology	31.7%	53.8%	−19.1%	57.6%
Act responsibly	10.4%	62.8%	−41.5%	57.9%
Defining economy	31.7%	57.2%	−20.2%	57.9%
Defining our coexistence	−2.7%	58.9%	−54.6%	47.2%
Control and guide our working life	14.8%	63.1%	−51.4%	54.2%
Create social division	3.3%	64.2%	−68.3%	42.2%
Control and guide our private life	−2.7%	67.0%	−76.0%	44.6%
Lead to isolation	2.7%	69.7%	−73.2%	49.7%
Make society more lazy	20.2%	70.5%	−60.7%	49.1%
Be influenced by a few	27.3%	54.9%	−59.0%	52.4%
Lead to less communication	23.5%	70.0%	−63.9%	52.2%
Be influenced by an elite	23.5%	65.1%	−65.6%	44.7%
Destroy more jobs	26.2%	66.3%	−64.5%	49.4%
Be hackable	60.7%	52.7%	−79.2%	42.1%

Items sorted from least to strongest discrepancy between likelihood and Evaluation. [a]Measured on two 4-point Likert scales and rescaled to −100% to +100%. Negative values indicate that the development is seen as unlikely respectively a negative assessment and positive values indicate a high estimated likelihood respectively positive evaluation.

TABLE 3 Correlations between AI assessment and the (explanatory) user factors.

Variable	Valence	Likelihood
1. Age in years	0.06	−0.10
2. Gender	−0.08	−0.15
3. Interpers. Trust	−0.10	0.18
4. Attitiude in technology interaction	−0.05	0.31
5. Trust in AI	−0.25	0.20
6. Distrust in AI	.22	−0.34
7. Average valence	–	−0.07
8. Average likelihood	−0.07	–

Note that gender is dummy-coded (0 = male, 1 = female).

features and consequences of AI seem desirable, but there is a lack of conviction that this will happen in such a positive way. Future research should further differentiate the concept of trust in this context: On the one hand, trust that the technology is reliable and not harmful, and on the other hand, trust that the technology can deliver what is promised to oneself or by others, i.e., trust as opposed to confidence.

In our explanatory analysis, we examined whether the expected likelihood was related to the valuation. However, this relationship was not confirmed, although the (non-significant) correlation was quite large. We refrain from making a final assessment and suggest that the correlation between valence and expected likelihood of occurrence should be re-examined with a larger sample and a more precise measurement of the target dimensions. This would provide a deeper understanding of whether there is a systematic bias between these two dimensions and at the same time allow, if possible, to derive distinguishable expectation profiles to compare user characteristics, e.g., between groups that are rather pessimistic about AI development, groups that have exaggerated expectations or naive ideas about the possibilities of AI, or groups that reject AI but fear that it will nevertheless permeate life.

5.1. Implications

As discussed above, AI is at the center of attention when it comes to innovative and "new" technology. Huge promises have been made about the impact, both positive and negative, that AI could have on society as a whole, but also on an individual level (Brynjolfsson and Mitchell, 2017; Ikkatai et al., 2022). This development has led to a shift in public attention and attitudes toward AI. Therefore, it is necessary to elaborate on this perception and attitude in order to find future research directions and possible educational approaches to increase people's literacy about AI and AI-based technologies. This discussion should also include a discourse on ethical implications, i.e., possible moral principles that should guide the way we research and develop AI. These principles should include individual, organizational and societal values as well as expert judgements about the context in which AI is appropriate or not (Awad et al., 2018; Liehner et al., 2021).

Previous research on this approach shows that cynical distrust of AI, i.e., the attitude that AI cannot be trusted *per se*, is a different construct from the same kind of distrust toward humans (Bochniarz et al., 2022). This implies that although AI is thought to be close to the human mind—at least in some circumstances—it is not confused with the human traits of hostility or emotionality. Importantly, according to Kolasinka et al., people have different evaluations of AI depending on the context (Kolasinska et al., 2019): When asked in which field of AI research they would invest an unlimited amount of money, people chose the fields of medicine and cybersecurity. There seems to be an overlap between the context in which AI is placed and the level of trust required in that specific context. For example, most people are not necessarily experts in cybersecurity or medicine. However, because of the trust placed in an IT expert, a doctor or any other expert, people generally do not question the integrity of these experts. AI is a similar matter, as people do not usually attribute emotionality to it, but rather objectivity, so they tend to trust its accuracy and disregard its potential for error (Cismariu and Gherhes, 2019; Liu and Tao, 2022).

Despite the benefits of AI, an accurate knowledge of its potential and limitations is necessary for a balanced and useful use of AI-based technology (Hick and Ziefle, 2022). Therefore, educational programmes for the general public and non-experts in the field of AI seems appropriate to provide a tool with which people can evaluate for themselves the benefits and barriers of this technology (Olari and Romeike, 2021). More research is needed to find out what are the most important and essential aspects of such an educational programme, but the map presented here may be a suitable starting point starting point to identify crucial topics.

6. Limitations and future work

Of course, this study is not without its limitations. First, the sample of 122 participants is not representative for the whole population of our country or even across countries. We therefore recommend that this method be used with a larger, more diverse sample, including participants of all ages, from different educational backgrounds and, ideally, from different countries and cultures. Nevertheless, the results presented here have their own relevance: Despite the relatively homogeneous young and educated sample, certain misconceptions about AI became apparent and imbalances in estimated likelihood and valuation could be identified. These could either be an obstacle to future technology development and adoption and/or are aspects that require societal debate and possibly regulation.

Second, participants responded to a short item on each topic and we refrained from explaining each idea in more detail. As a result, the answers to these items may have been shaped by affective rather than cognitively considered considerations. However, this is not necessarily a disadvantage. On the one hand, this approach made it easier to explore a wide range of possible ways in which AI might affect our future. On the other hand, and more importantly, we as humans are not rational agents, but most of our decisions and behavior are influenced by cognitive biases and our affect (i.e., "affect heuristic") (Finucane et al., 2000; Slovic et al., 2002). In this respect, this study contributes to affective technology evaluation, which nonetheless influences evaluation and use.

From a methodological point of view, asking for ratings with only two single items leads to a high variance and makes it difficult to examine individual aspects in detail. Although this allowed us to address a variety of different issues, future work should select specific aspects and examine them in more detail. Consequently, future work may further integrate other concepts, such as the impact of AI on individual mobility, public safety, or even warfare through automation and control. However, the present approach allowed us to keep the survey reasonably short, which had a positive effect on response attention and unbiased dropout rates.

Finally, we propose the integration of expert judgement into this cartography. We suspect that there are considerable differences between expert and lay assessments, particularly in the assessment of the expected likelihood of the developments in question. Again, it is the differences between expert and lay expectations that are particularly relevant for informing researchers and policy makers.

7. Conclusion

The continuing and increasing pervasiveness of AI in almost all personal and professional contexts will reshape our future, how we interact with technology, and how we interact with others using technology. Responsible research and innovation on AI-based products and services requires us to balance technical advances and economic imperatives with individual, organizational, and societal values (Burget et al., 2017; Owen and Pansera, 2019).

This work suggests that the wide range of potential AI applications is assessed differently in terms of perceived likelihood and perceived valence as a measure of acceptability. The empirically derived criticality map makes this assessment visible and highlights issues with urgent potential for research, development, and governance and can thus contribute to responsible research and innovation of AI.

We also found individual differences in perceptions of AI that may threaten both people's ability to participate in societal debates about AI and to adequately adapt their future skill sets to compete with AI in the future of work. It is a political issue, not a technological one, in which areas AI can influence our lives and society, and to what extent. As a society, we need to discuss and debate the possibilities and limits of AI in a wide range of applications and define appropriate regulatory frameworks. For this to happen, we all need to have a basic understanding of AI so that we can participate in a democratic debate about its potential and its limits. Free online courses for adults such as "Elements of AI" and modern school curricula that teach the basics of digitalisation and AI are essential for this (Olari and Romeike, 2021; Marx et al., 2022).

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://osf.io/f9ek6/.

Ethics statement

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. The patients/participants provided their written informed consent to participate in this study.

Author contributions

PB and RP designed the study. PB wrote the original draft of the manuscript and coordinated the analysis and writing, while RP and AH made substantial contributions to the motivation, related work, and discussion sections of the manuscript. MZ supervised the work and acquired the funding for this research. All authors contributed to the article and approved the submitted version.

Funding

This work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy—EXC-2023 Internet of Production—390621612 and the VisuAAL project "Privacy-Aware and Acceptable Video-Based Technologies and Services for Active and Assisted Living", funded by the European Union's Horizon 2020 Research and Innovation Programme under the Marie Skłodowska-Curie Grant Agreement No. 861091.

Acknowledgments

We would particularly like to thank Isabell Busch for her support in recruiting participants, Mohamed Behery for his encouragement in writing this manuscript, Luca Liehner for feedback on the key figure, and Johannes Nakayama and Tim Schmeckel for their support in analysis and writing. Thank you to the referees and the editor for their valuable and constructive input. The writing of this article was partly supported by ChatGPT, DeepL, and DeepL Write.

Conflict of interest

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

Publisher's note

References

Acemoglu, D., and Restrepo, P. (2017). The Race Between Machine and Man. *Am. Econ. Rev.* 108, 1488–1542. doi: 10.3386/w22252

Almars, A. M., Gad, I., and Atlam, E.-S. (2022). "Applications of AI and IoT in COVID-19 vaccine and its impact on social life," in *Medical Informatics and Bioimaging Using Artificial Intelligence* (Springer), 115–127.

Araujo, T., Helberger, N., Kruikemeier, S., and Vreese, C. H. de (2020). In AI we trust? Perceptions About Automated Decision-making by Artificial Intelligence. *AI Society* 35, 611–623. doi: 10.1007/s00146-019-00931-w

Awad, E., Dsouza, S., Kim, R., Schulz, J., Henrich, J., Shariff, A., et al. (2018). The moral machine experiment. *Nature* 563, 59–64. doi: 10.1038/s41586-018-0637-6

Beierlein, K., C. (2014). Interpersonales vertrauen (KUSIV3). *Zusammenstellung sozialwissenschaftlicher Items und Skalen (ZIS)*. doi: 10.6102/zis37

Binz, M., and Schulz, E. (2023). Using cognitive psychology to understand GPT-3. *Proc. Natl. Acad. Sci. U.S.A.* 120, e2218523120. doi: 10.1073/pnas.2218523120

Bochniarz, K. T., Czerwiński, S. K., Sawicki, A., and Atroszko, P. A. (2022). Attitudes to AI among high school students: Understanding distrust towards humans will not help us understand distrust towards AI. *Pers. Ind. Diff.* 185, 111299. doi: 10.1016/j.paid.2021.11129

Brauner, P., Dalibor, M., Jarke, M., Kunze, I., Koren, I., Lakemeyer, G., et al. (2022). A computer science perspective on digital transformation in production. *ACM Trans. Internet Things* 3, 1–32. doi: 10.1145/3502265

Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., et al. (2020). "Language models are few-shot learners," in *Advances in neural information processing systems*, eds H. Larochelle, M. Ranzato, R. Hadsell, M. F. Balcan, and H. Lin (Curran Associates, Inc.), 1877–1901. Available online at: https://proceedings.neurips.cc/paper/2020/file/1457c0d6bfcb4967418bfb8ac142f64a-Paper.pdf

Brynjolfsson, E., and Mitchell, T. (2017). What can machine learning do? Workforce implications. *Science* 358, 1530–1534. doi: 10.1126/science.aap8062

Bughin, J., Seong, J., Manyika, J., Chui, M., and Joshi, R. (2018). Notes from the AI frontier: modeling the impact of AI on the world economy. *McKinsey Glob. Inst.* 4.

Burbach, L., Halbach, P., Plettenberg, N., Nakayama, J., Ziefle, M., and Calero Valdez, A. (2019). ""Hey, Siri", "Ok, Google", "Alexa". Acceptance-relevant factors of virtual voice-assistants," in *2019 IEEE International Professional Communication Conference (ProComm)* (Aachen: IEEE), 101–111.

Burget, M., Bardone, E., and Pedaste, M. (2017). Definitions and conceptual dimensions of responsible research and innovation: a literature review. *Sci. Eng. Ethics* 23, 1–19. doi: 10.1007/s11948-016-9782-1

Cambridge Dictionary (2022). Cambridge dictionary. *Artificial Intelligence*. Available online at: https://dictionary.cambridge.org/dictionary/english/artificial-intelligence (accessed December 1, 2022).

Cismariu, L., and Gherhes, V. (2019). Artificial intelligence, between opportunity and challenge. *Brain* 10, 40–55. doi: 10.18662/brain/04

Collingridge, D. (1982). *Social Control of Technology*. Continuum International Publishing Group Ltd.

Corea, F. (2019). "AI knowledge map: How to classify AI technologies," in *An Introduction to Data. Studies in Big Data*, Vol 50 (Cham: Springer). doi: 10.1007/978-3-030-04468-8_4

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *psychometrika* 16, 297–334.

Cugurullo, F., and Acheampong, R. A. (2023). Fear of AI: An inquiry into the adoption of autonomous cars in spite of fear, and a theoretical framework for the study of artificial intelligence technology acceptance. *AI & Society*, 1–16. doi: 10.1007/s00146-022-01598-6

Dalkey, N., and Helmer, O. (1963). An experimental application of the delphi method to the use of experts. *Manag. Sci.* 9, 458–467. doi: 10.1287/mnsc.9.3.458

Dienes, Z. (2008). *Understanding Psychology as a Science-An Introduction to Scientific and Statistical Inference, 1st Edn* (London: Red Globe Press).

Finucane, M. L., Alhakami, A., Slovic, P., and Johnson, S. M. (2000). The affect heuristic in judgments of risks and benefits. *J. Behav. Decis. Making* 13, 1–17. doi: 10.1002/(SICI)1099-0771(200001/03)13:1<::AID-BDM333>.0.CO;2-S

Floridi, L., and Cowls, J. (2022). "A unified framework of five principles for AI in society," in *Machine Learning and the City: Applications in Architecture and Urban Design*, 535–545.

Flowers, J. C. (2019). "Strong and weak AI: deweyan considerations," in *AAAI Spring Symposium: Towards Conscious AI Systems*.

Foot, P. (1967). *The Problem of Abortion and the Doctrine of the Double Effect*. Oxford: Oxford Review.

Fosso Wamba, S., Bawack, R. E., Guthrie, C., Queiroz, M. M., and Carillo, K. D. A. (2021). Are we preparing for a good AI society? A bibliometric review and research agenda. *Technol. Forecast. Soc. Change* 164, 120482. doi: 10.1016/j.techfore.2020.120482

Franke, T., Attig, C., and Wessel, D. (2019). A personal resource for technology interaction: development and validation of the affinity for technology interaction (ATI) scale. *Int. J. Human Comput. Interact.* 35, 456–467. doi: 10.1080/10447318.2018.1456150

Gibson, R. (2019). *Desire in the Age of Robots and AI: An Investigation in Science Fiction and Fact*. Springer.

Grace, K., Salvatier, J., Dafoe, A., Zhang, B., and Evans, O. (2018). When will AI exceed human performance? Evidence from AI experts. *J. Artif. Intell. Res.* 62, 729–754. doi: 10.1613/jair.1.11222

Gunkel, D. J. (2012). *The Machine Question: Critical Perspectives on AI, Robots, and Ethics*. Cambridge, MA: MIT Press.

Harari, Y. N. (2017). *Homo Deus: A Brief History of Tomorrow*. Palatine, IL: Harper.

Hick, A., and Ziefle, M. (2022). "A qualitative approach to the public perception of AI," in *IJCI Conference Proceedings,* eds D. C. Wyld et al., 01–17.

Hirsch-Kreinsen, H. (2023). Artificial intelligence: A "promising technology." *AI & Society* 2023, 1–12. doi: 10.1007/s00146-023-01629-w

Hoff, K. A., and Bashir, M. (2015). Trust in automation: integrating empirical evidence on factors that influence trust. *Human Factors* 57, 407–434. doi: 10.1177/0018720814547570

Ikkatai, Y., Hartwig, T., Takanashi, N., and M., Y. H. (2022). Segmentation of ethics, legal, and social issues (ELSI) related to AI in Japan, the united states, and Germany. *AI Ethics* doi: 10.1007/s43681-022-00207-y

Jovanovic, M., Mitrov, G., Zdravevski, E., Lameski, P., Colantonio, S., Kampel, M., et al. (2022). Ambient assisted living: Scoping review of artificial intelligence models, domains, technology, and concerns. *J. Med. Internet Res.* 24, e36553. doi: 10.2196/36553

Jumalon, G. (2022). TL;DR–someone entered an art competition with an AI-generated piece and won the first prize. Available online at: https://twitter.com/GenelJumalon/status/1564651635602853889 (accessed November 28, 2022).

Kelly, S., Kaye, S.-A., and Oviedo-Trespalacios, O. (2023). What factors contribute to the acceptance of artificial intelligence? A systematic review. *Telematics Inf.* 77, 101925. doi: 10.1016/j.tele.2022.101925

Klos, A., Rosenbaum, M., and Schiffmann, W. (2020). "Emergency landing field identification based on a hierarchical ensemble transfer learning model," in *IEEE 8th international symposium on computing and networking (CANDAR)* (Naha: IEEE), 49–58.

Kolasinska, A., Lauriola, I., and Quadrio, G. (2019). "Do people believe in artificial intelligence? A cross-topic multicultural study," in *Proceedings of the 5th EAI International Conference on Smart Objects and Technologies for social good GoodTechs '19.* (New York, NY: Association for Computing Machinery), 31–36.

Kulida, E., and Lebedev, V. (2020). "About the use of artificial intelligence methods in aviation," in *13th International Conference on Management of Large-Scale System Development (MLSD)*, 1–5. doi: 10.1109/MLSD49919.2020.9247822

Kulkarni, S., Seneviratne, N., Baig, M. S., and Khan, A. H. A. (2020). Artificial intelligence in medicine: where are we now? *Acad. Radiol.* 27, 62–70. doi: 10.1016/j.acra.2019.10.001

Lecun, Y., Bengio, Y., and Hinton, G. (2015). Deep Learning. *Nature* 521, 436–444. doi: 10.1038/nature14539

Liehner, G. L., Brauner, P., Schaar, A. K., and Ziefle, M. (2021). Delegation of moral tasks to automated agents The impact of risk and context on trusting a machine to perform a task. *IEEE Trans. Technol. Soc.* 3, 46–57. doi: 10.1109/TTS.2021.3118355

Lin, H.-Y. (2023). Standing on the shoulders of AI giants. *Computer* 56, 97–101. doi: 10.1109/MC.2022.3218176

Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., et al. (2017). A survey on deep learning in medical image analysis. *Med. Image Anal.* 42, 60–88. doi: 10.1016/j.media.2017.07.005

Liu, K., and Tao, D. (2022). The roles of trust, personalization, loss of privacy, and anthropomorphism in public acceptance of smart healthcare services. *Comput. Human Behav.* 127, 107026. doi: 10.1016/j.chb.2021.107026

Liu, R., Rizzo, S., Whipple, S., Pal, N., Pineda, A. L., Lu, M., et al. (2021). Evaluating eligibility criteria of oncology trials using real-world data and AI. *Nature* 592, 629–633. doi: 10.1038/s41586-021-03430-5

Makridakis, S. (2017). The forthcoming artificial intelligence (AI) revolution: its impact on society and firms. *Futures* 90, 46–60. doi: 10.1016/j.futures.2017.03.006

Marcus, G., and Davis, E. (2019). *Rebooting AI: Building Artificial Intelligence We Can Trust*. New York, NY: Pantheon Books.

Marx, E., Leonhardt, T., Baberowski, D., and Bergner, N. (2022). "Using matchboxes to teach the basics of machine learning: An analysis of (possible) misconceptions," in *Proceedings of the Second Teaching Machine Learning and Artificial Intelligence Workshop Proceedings of Machine Learning Research*, eds K. M. Kinnaird, P. Steinbach, and O. Guhr (PMLR), 25–29. Available online at: https://proceedings.mlr.press/v170/marx22a.html

Mayer, R. C., Davis, J. H., and Schoorman, F. D. (1995). An integrative model of organizational trust. *Acad. Manag. Rev.* 20, 709–734.

Mc Knight, D. H., Choudury, V., and Kacmar, C. (2002). Developing and validating trust measure for e-commerce: an integrative typology. *Inf. Syst. Res.* 13, 334–359. doi: 10.1287/isre.13.3.334.81

McCarthy, J., Minsky, M. L., Rochester, N., and Shannon, C. E. (2006). A proposal for the dartmouth summer research project on artificial intelligence (August 31, 1955). *AI Mag.* 27, 12–12. doi: 10.1609/aimag.v27i4.1904

McCrae, R. R., and Costa, P. T. (1987). Validation of the five-factor model of personality across instruments and observers. *J. Pers. Soc. Psychol.* 52, 81.

Oden, L., and Witt, T. (2020). "Fall-detection on a wearable micro controller using machine learning algorithms," in *IEEE International Conference on Smart Computing 2020 (SMARTCOMP)* (Bologna: IEEE), 296–301.

Oksanen, A., Savela, N., Latikka, R., and Koivula, A. (2020). Trust toward robots and artificial intelligence: an experimental approach to human–technology interactions online. *Front. Psychol.* 11, 568256. doi: 10.3389/fpsyg.2020.568256

Olari, V., and Romeike, R. (2021). "Addressing AI and data literacy in teacher education: a review of existing educational frameworks," in *The 16th Workshop in Primary and Secondary Computing Education WiPSCE '21* (New York, NY: Association for Computing Machinery).

Onnasch, L., and Roesler, E. (2020). A taxonomy to structure and analyze human–robot interaction. *Int. J. Soc. Rob.* 13, 833–849. doi: 10.1007/s12369-020-00666-5

Owen, R., Macnaghten, P., and Stilgoe, J. (2012). Responsible research and innovation: from science in society to science for society, with society. *Sci. Public Policy* 39, 751–760. doi: 10.1093/scipol/scs093

Owen, R., and Pansera, M. (2019). Responsible innovation and responsible research and innovation. *Handbook Sci. Public Policy*, 26–48. doi: 10.4337/9781784715946.00010

Philipsen, R., Brauner, P., Biermann, H., and Ziefle, M. (2022). I am what i am–roles for artificial intelligence from the users' perspective. *Artif. Intell. Soc. Comput.* 28, 108–118. doi: 10.54941/ahfe1001453

Rao, Q., and Frtunikj, J. (2018). "Deep learning for self-driving cars: Chances and challenges," in *Proceedings of the 1st International Workshop on Software Engineering for AI in Autonomous Systems*, 35–38. doi: 10.1145/3194085.3194087

Rashidi, P., and Mihailidis, A. (2013). A survey on ambient-assisted living tools for older adults. *IEEE J. Biomed. Health Inform.* 17, 579–590. doi: 10.1109/JBHI.2012.2234129

Reeves, B., and Nass, C. (1996). *The Media Equation–How People Treat Computers, Television, and New Media Like Real People and Places*. Cambridge: Cambridge University Press.

Robb, D. A., Ahmad, M. I., Tiseo, C., Aracri, S., McConnell, A. C., Page, V., et al. (2020). "Robots in the danger zone: Exploring public perception through engagement," in *Proceedings of the 2020 ACM/IEEE International Conference on Human-Robot Interaction HRI '20.* (New York, NY: Association for Computing Machinery), 93–102.

Rodriguez-Bustelo, C., Batista-Foguet, J. M., and Serlavós, R. (2020). Debating the future of work: The perception and reaction of the spanish workforce to digitization and automation technologies. *Front. Psychol.* 11, 1965. doi: 10.3389/fpsyg.2020.01965

Russell, S., and Norvig, P. (2009). *Artificial Intelligence: A Modern Approach, 3rd Edn*. Hoboken, NJ: Prentice Hall Press.

Sindermann, C., Sha, P., Zhou, M., Wernicke, J., Schmitt, H. S., Li, M., et al. (2021). Assessing the attitude towards artificial intelligence: introduction of a short measure in german, chinese, and english language. *KI-Künstliche Intelligenz* 35, 109–118. doi: 10.1007/s13218-020-00689-0

Slovic, P. (1987). Perception of risk. *Science* 236, 280–285. doi: 10.1126/science.3563507

Slovic, P., Finucane, M., Peters, E., and MacGregor, D. G. (2002). "The affect heuristic," in *Heuristics and Biases: The Psychology of Intuitive Judgment*, eds T. Gilovich, D. Griffin, and D. Kahneman (Cambridge: Cambridge University Press), 397–420.

Smith, A., and Anderson, J. (2014). AI, robotics, and the future of jobs. *Pew Res. Center* 6, 51. Available online at: https://www.pewresearch.org/internet/2014/08/06/future-of-jobs/

Statista (2022). Artificial intelligence (AI) worldwide-statistics & facts. Available online at: https://www.statista.com/topics/3104/artificial-intelligence-ai-worldwide/#dossier-chapter1 (accessed November 28, 2022)

Strich, F., Mayer, A.-S., and Fiedler, M. (2021). What do i do in a world of artificial intelligence? Investigating the impact of substitutive decision-making AI systems on employees' professional role identity. *J. Assoc. Inf. Syst.* 22, 9. doi: 10.17705/1jais.00663

Taherdoost, H. (2018). A review of technology acceptance and adoption models and theories. *Procedia Manufact.* 22, 960–967. doi: 10.1016/j.promfg.2018.03.137

Topol, E. (2019). *Deep Medicine: How Artificial Intelligence Can Make Healthcare Human Again*. London: Hachette UK.

Vaishya, R., Javaid, M., Khan, I. H., and Haleem, A. (2020). Artificial intelligence (AI) applications for COVID-19 pandemic. *Diabetes Metabolic Syndrome Clin. Res. Rev.* 14, 337–339. doi: 10.1016/j.dsx.2020.04.012

West, D. M. (2018). *The Future of Work: Robots, AI, and Automation*. Washington, DC: Brookings Institution Press.

Wilkowska, W., Brauner, P., and Ziefle, M. (2018). "Rethinking technology development for older adults: a responsible research and innovation duty," in *Aging, Technology and Health*, eds R. Pak and A. C. McLaughlin (Cambridge, MA: Academic Press), 1–30.

Wolff, J., Pauling, J., Keck, A., and Baumbach, J. (2020). The economic impact of artificial intelligence in health care: Systematic review. *J. Med. Internet Res.* 22, e16866. doi: 10.2196/16866

Young, A. T., Amara, D., Bhattacharya, A., and Wei, M. L. (2021). Patient and general public attitudes towards clinical artificial intelligence: a mixed methods systematic review. *Lancet Digital Health* 3, e599–e611. doi: 10.1016/S2589-7500(21)00132-1

Zuiderwijk, A., Chen, Y.-C., and Salem, F. (2021). Implications of the use of artificial intelligence in public governance: a systematic literature review and a research agenda. *Government Inform. Q.* 38, 101577. doi: 10.1016/j.giq.2021.101577

frontiers | Frontiers in Computer Science

TYPE Original Research
PUBLISHED 31 March 2023
DOI 10.3389/fcomp.2023.1129340

OPEN ACCESS

EDITED BY
Lynsay Shepherd,
Abertay University, United Kingdom

REVIEWED BY
Curtis Gittens,
The University of the West Indies, Cave Hill, Barbados
Pei-Chun Lin,
Feng Chia University, Taiwan

*CORRESPONDENCE
Elie Maalouly
✉ maalouly.elie@irl.sys.es.osaka-u.ac.jp
Ryuji Yamazaki
✉ yamazaki@irl.sys.es.osaka-u.ac.jp

SPECIALTY SECTION
This article was submitted to Human-Media Interaction, a section of the journal Frontiers in Computer Science

RECEIVED 21 December 2022
ACCEPTED 14 March 2023
PUBLISHED 31 March 2023

CITATION
Maalouly E, Yamazaki R, Nishio S, Nørskov M, Kamaga K, Komai S, Chiba K, Atsumi K and Akao K-I (2023) Assessing the effect of dialogue on altruism toward future generations: A preliminary study.
Front. Comput. Sci. 5:1129340.
doi: 10.3389/fcomp.2023.1129340

Assessing the effect of dialogue on altruism toward future generations: A preliminary study

Elie Maalouly[1]*, Ryuji Yamazaki[2]*, Shuichi Nishio[2], Marco Nørskov[3], Kohei Kamaga[4], Shoji Komai[5], Kiyoshi Chiba[6], Keiichiro Atsumi[6] and Ken-Ichi Akao[6]

[1]Graduate School of Engineering Science, Osaka University, Osaka, Japan, [2]Institute for Open and Transdisciplinary Research Initiatives, Osaka University, Osaka, Japan, [3]Department for Philosophy and The History of Ideas, Aarhus University, Aarhus, Denmark, [4]Faculty of Economics, Sophia University, Tokyo, Japan, [5]Faculty of Engineering, International Professional University of Technology in Tokyo, Tokyo, Japan, [6]School of Social Sciences, Waseda University, Tokyo, Japan

Introduction: Despite the abundance of evidence on climate change and its consequences on future generations, people, in general, are still reluctant to change their actions and behaviors toward the environment that would particularly benefit posterity. In this study, we took a preliminary step in a new research direction to explore humans' altruistic behavior toward future generations of people and whether it can be affected by dialogue.

Methods: We used an android robot called Telenoid as a representative of future generations by explaining that the robot is controlled by an Artificial Intelligence (AI) living in a simulation of our world in the future. To measure people's altruistic behavior toward it, we asked the participants to play a round of the Dictator Game with the Telenoid before having an interactive conversation with the Telenoid and then playing another round.

Results: On average, participants gave more money to the Telenoid in the second round (after having an interactive conversation). The average amount of money increased from 20% in the first to about 30% in the second round.

Discussion: The results indicate that the conversation with the robot might have been responsible for the change in altruistic behavior toward the Telenoid. Contrary to our expectations, the personality of the participants did not appear to have an influence on their change of behavior, but other factors might have contributed. We finally discuss the influence of other possible factors such as empathy and the appearance of the robot. However, the preliminary nature of this study should deter us from making any definitive conclusions, but the results are promising for establishing the ground for future experiments.

KEYWORDS

altruism, android robot, dialogue, dictator game, future generations

1. Introduction

Climate change is a major threat to our planet and the human race. But it is a future issue that might not affect our current generation. However, if there is such a threat to future generations of humans, then why are we so indifferent and passive? Do we lack altruism toward future generations of people? And if so, how can we influence people's altruistic behavior to posterity?

It is expected that climate change would have a wide range of effects, including changes to ecosystems (Leemans and Eickhout, 2004), effects on human systems including water resources (Arnell and Lloyd-Hughes, 2014), forced human migration (Barnett and Adger, 2003), and extensive ocean acidification (Caldeira and Wickett, 2003). However, despite the abundance of evidence available on climate change, people are still hesitant to alter their energy-related choices and actions. Efforts to persuade people to adopt environmentally friendly habits have often been somewhat unsuccessful (Whitmarsh and O'Neill, 2010; Chu and Majumdar, 2012). Therefore, understanding how to influence this altruistic behavior can be crucial for the preservation of our environment and species, whereby carelessness and negligence might bring on devastating consequences.

But there is more than the preservation of the environment, as engaging in altruistic behavior rewards the person with multiple health benefits. Altruism is beneficial for a person's emotional health and can dramatically improve their peace of mind (Brunier et al., 2002), in addition to improving confidence, self-esteem, self-awareness, daily function, and reducing depression (Uccelli et al., 2004).

We aim to investigate whether humans are capable of altruistic behavior toward future generations of people and what factors might have an effect on this behavior. In this study, we would like to explore dialogue specifically as a factor in affecting altruistic behavior toward future generations. The main questions we seek answers to are whether humans can feel compassion for a complete stranger from a different timeline and how dialogue might affect it. In the following section, we will briefly introduce and define our understanding of altruistic behavior and empathy. We will then discuss how altruistic behavior is measured in experiments and how empathy, dialogue, and personality might influence it. We will begin by discussing how altruism is measured in experiments, namely using the Dictator Game.

2. Related work

2.1. Defining altruism

There is a big disagreement among researchers on how to define altruism and how it is linked to prosocial behavior. In more than 74% of the articles published that include the keywords "prosocial behavior," "prosocial lies," "altruism" and "altruistic behavior," there was no relevant definition included in the article (Pfattheicher et al., 2022). Therefore, properly defining these constructs is important to communicate how we understand altruistic behavior in this study.

The fact that prosocial and altruistic behaviors indicate a "positive" social behavior toward one or several others is common in nearly all definitions. These behaviors are most frequently defined as promoting or having as their primary goal the promotion of the welfare of others. Where the definitions differ is mainly in what constitutes prosocial or altruistic behavior. With regard to the final point, there are three perspectives on what is considered prosocial behavior, and altruism is defined differently for each of these perspectives. First, is the intentionalist perspective, where the definition of prosocial behavior emphasizes the intentional nature of the behavior. According to Batson and Powell (2003), "prosocial behavior covers the broad range of actions intended to benefit one or more people other than oneself". The second perspective is the consequential perspective, where the definition of prosocial behavior emphasizes the consequences of the act instead of the intentions and motives. According to Schroeder and Graziano (2015), prosocial behavior is defined as "any action that benefits another." And finally, there is the societal perspective that emphasizes societal approval, where prosocial behavior is simply defined as "behavior that is valued by the individual's society" (Dovidio, 1984).

In this study, we will be working within an intentionalist perspective of prosocial behavior. From an intentionalist perspective, altruism can be defined as a motivational state or a subtype of prosocial behavior. For example, Batson (2010) defines altruism as "motivation with the ultimate goal of increasing another's welfare." Similarly, Eisenberg and Miller (1987) defines altruism as "prosocial behavior which is not performed with the expectation of receiving external rewards or avoiding externally produced aversive stimuli or punishments." Therefore according to the intentionalist perspective, prosocial behavior and altruism are two separate but related ideas. While prosocial behavior refers only to the deed itself, altruism also refers to the motivation behind the action.

Going forward, we will rely on Batson's definition of altruism which is, "motivation with the ultimate goal of increasing another's welfare."

2.2. Empathy and altruism

Now that we have confined our definition of altruistic behavior to Batson's, it is important now to understand what causes such behavior. According to Batson, the empathy-altruism hypothesis is the main cause of altruistic behavior (Schroeder and Graziano, 2015). According to the empathy-altruism hypothesis, altruistic motivation is produced by empathic concern (Batson, 1987, 2011). Although empathy and empathic concern are often used interchangeably, here it is again important to define what is meant by empathy and "empathic concern." Empathic concern is the "other-oriented emotional response elicited by and congruent with the perceived welfare of a person in need" (Batson, 2010). Instead of feeling as the other person does, this emotion entails feeling for the other. Although this emotion has been called different names in different fields, such as sympathy, tenderness, compassion, pity, and empathy (Stotland, 1969; Hoffman, 1975; Krebs, 1975), the term empathic concern is used to emphasize that the emotion is due to another's suffering compared to the term empathy is much broader.

Let us unpack the meaning of the empathic concern, stated previously as the "other-oriented emotional response elicited by and congruent with the perceived welfare of a person in need."

1. The term "congruent" is a reference to the valence instead of the content of the emotion, meaning negative if the perceived welfare of the other is negative.
2. Even though empathy is defined very broadly to encompass (for example) experiencing empathy for another person's success, only empathy experienced while someone is thought to be in need is believed to induce altruistic motivation.

3. Empathetic concern includes a vast array of feelings rather than a single, distinct emotion. It includes sympathy, tenderness, sorrow, compassion, sadness, softheartedness, upset, concern, grief, and distress.
4. Empathic concern is "other-oriented" in that it entails feeling for the other.

The term "empathy" alone is very broad and has been applied to mean different things by different researchers which makes it thus essential to properly define. It has been used to mean:

- Knowing the feelings and thoughts of another (also known as "cognitive empathy").
- Adopting another's stance or neurological reaction (also known as "physiological sympathy" or "facial empathy").
- Assuming another person's perspective (also known as "affective empathy").
- Experiencing distress at seeing the suffering of another (also known as "empathic distress").
- Putting oneself in another's shoes to consider and feel what they would (also known as "perspective taking").

The other-oriented feeling that we call empathic concern is different from all of the above meanings of empathy. Therefore, in this study, we define the term empathy to refer to empathic concern.

There is a great deal of evidence to back up the claim that empathy drives altruistic behavior. Aronfreed (1968) and Aderman and Berkowitz (1970) presented the first supporting data, where they each designed experimental conditions intended to promote or prevent empathy for a person in need. Each study found that helping increased in the experimental settings intended to promote empathy, and both pairs of authors came to the conclusion that greater empathy was associated with greater helping. These findings were supported by a great number of experiments, some of which are (Harris and Huang, 1973; Krebs, 1975; Coke et al., 1978; Dovidio et al., 1990). For an extensive review of the experiments done to support this hypothesis, please see Batson (2011). Reports of over 30 experiments testing the empathy-altruism hypothesis, and the results so far have supported the hypothesis.

The empathy-altruism hypothesis states that altruistic motivation is produced by empathy and this seems to be heavily supported by experimental evidence, but what causes empathy in the first place? Generally, perceiving the other as in need and valuing the other's welfare appear to be the two conditions that must be met in order to experience empathy in daily life (Batson, 2011).

For perceiving need, one must be able to distinguish between the other's current situation and their ideal condition on one or more well-being dimensions (Latané and Darley, 1970; Clark and Word, 1972, 1974). The dimensions involved are negative affect, the absence of bodily discomfort, anxiety, danger, stress, and disease, in addition to positive affect, the presence of physical pleasure, security, and satisfaction. There has been evidence to support the hypothesis that empathy requires the perception of need. Berger (1962) asked participants to watch a target person complete a task. He misled these participants into thinking that the target either experienced electric shocks (electric-shock condition) or not (no-shock condition). The target also moved his arm in response to shock (a movement requirement) or not (no-movement condition). Everyone who took part in the study was informed that they would not experience any shocks during the experiment. Berger argued, first, that for an observer to assume that the target was in pain (i.e., need), both a painful stimulus (shock) and a distress reaction (movement) were required. Second, he reasoned that participants in his experiment should only have a physiological response when they inferred that the target was in pain if they were feeling empathy for the target rather than dread or worry about the shock itself. Berger thus expected that only those who participated in the shock/movement condition would exhibit elevated physiological arousal as only they would draw this conclusion. There was some information lacking for participants in each of the other three circumstances that would have allowed one to infer pain. Results conformed to expectations. Participants in the shock/movement condition were more physiologically aroused when seeing the target than were participants in the other three situations, which is consistent with the hypothesis that people can experience empathy when witnessing another perceived to be in need. This conclusion was further supported by subsequent studies (Bandura and Rosenthal, 1966; Craig and Lowery, 1969). Although the investigations just mentioned show that people respond physiologically to the perception of need in another, Stotland (1969) showed in a series of experiments that this physiological response represents empathy for others and that empathy can be increased with perspective taking.

To perceive another as in need is not enough. As mentioned earlier, the second condition required for experiencing empathy is valuing the other's welfare. We are less inclined to consider how a person is impacted by a need if we don't value the welfare of the person we perceive to be in need. If we dislike someone, it is common to place a negative value on their welfare. In that case, perceiving that they are in need might result in pleasure at witnessing their suffering as opposed to feelings of empathy (Zillman and Cantor, 1977; Lanzetta and Englis, 1989; Singer et al., 2006). Alternatively, placing a positive value on another's welfare makes it highly likely to adopt an other-oriented value appraisal of these events and take into account how this individual is affected by the events in his or her life (Batson, 2011). As long as there are no obvious reasons for antipathy, people generally tend to instinctively place a positive value (or at least a moderate value) on the welfare of others, even complete strangers (Batson, 2011).

2.3. Dialogue and empathy

Now we can look into how dialogue might affect empathy. While few studies have looked into the relationship between dialogue and empathy, there have been a few key findings that support it. Conversations, according to Nishida (2012), are made up of both verbal and nonverbal communication. Communication in a dialogue is facilitated by joint actions and words shared by two individuals, such as partaking in comparable behaviors and using similar language. Gould and MacNeil Gautreau (2014) looked into

the relationship between empathy and conversational enjoyment. Their findings revealed a link between older individuals' stated levels of satisfaction in their interactions and their level of empathetic care for others. Depending on the person conversing, the topic, length, and enjoyment of the conversation can differ. While there is evidence of a link between empathic concern and conversational enjoyment (Gould and MacNeil Gautreau, 2014), little research has been done on the types of conversations that can build empathy. Self-disclosure, on the other hand, has been widely utilized as a concept to evaluate the increase in intimacy (Laurenceau et al., 2004). Disclosure could be regarded as factual and descriptive or emotional and evaluative (Morton, 1978). Emotional disclosure entails more feelings than factual disclosure, which is more impersonal and insignificant (Morton, 1978). There may be a link between emotional disclosure and the level of intimacy experienced between people (Kruglanski and Higgins, 2013). While disclosure isn't required in every conversation, these findings imply a link between the sort of disclosure utilized and increased empathy. Finally, Andreoni and Rao (2011) showed that two-way verbal communication between participants of a Dictator Game elicited a higher altruistic behavior from allocators compared to one-way communication and no communication at all.

2.4. Altruism in experiments

Experiments on altruism create significant research challenges. Investigating altruism in experiments mainly centers around removing any viable ulterior motive ingrained in selfishness (Andreoni et al., 2010). The Dictator Game (Forsythe et al., 1994) is commonly used as a measure of altruism because it gives a simple and relatively pure assessment of altruistic vs. self-interested behavior and is frequently referred to as a measure of unconditional kindness (Ben-Ner and Halldorsson, 2010). This line of research started with the Ultimatum Game (Güth et al., 1982), where a proposer proposes a split to be divided between him and a responder. The split goes on if the responder accepts, whereas if he rejects both parties get nothing. In the Ultimatum Game, people appear to behave in a non-optimal way, where proposers give out fair deals. But is this altruism or just a fear of rejection? In other words, the strategy involved in Ultimatum Game might force a selfish proposer to give a fair split. To solve this, Forsythe et al. (1994) proposed the Dictator Game where the responder has no say in carrying out the split, The proposers are free to decide on any split they want. In the Dictator Game, preserving the entire endowment is the optimal decision, although a considerable portion of participants gives money away. According to a meta-analysis across 131 papers by Engel (2011), dictators gave on average 28.35% of their endowment, with a bimodal underlying distribution with peaks at 0% and 50%.

Although the Dictator Game is a very common measure of altruistic behavior in experiments, there might be other potential motives involved in the Dictator Game that are not purely altruistic, such as demand characteristics (Bardsley, 2008) and self-signaling (Grossman, 2015).

2.5. Dictator game and personality

Several studies have linked personality traits to the Dictator Game altruism using the five-factor model for personality. The outcomes, however, were mixed. None of the traditional five factors consistently predicted altruism in the Dictator Game. Even agreeableness, which is generally believed to be the one trait out of the five that should positively predict altruistic behavior, has only been proven to predict Dictator Game altruism positively in certain studies (Ben-Ner et al., 2008; Becker et al., 2012; Baumert et al., 2014), but not (or even negatively) in others (Weitzel et al., 2010; Ben-Ner and Kramer, 2011; Visser and Roelofs, 2011). The effect of personality thus seems inconclusive and might be worth pursuing.

2.6. Approach and hypothesis

We believe dialogue can be a powerful factor in promoting altruistic behavior mainly due to its power on building empathy. Even though there hasn't been a lot of work done on the direct relationship between dialogue and empathy as we have discussed in Section 2.3, we believe dialogue can be an important tool for promoting empathy. As we are working within Batson's definition of altruism and empathy, and within the empathy-altruism hypothesis that specifies empathy as a primary cause of altruistic behavior, we believe dialogue might be a great tool for promoting empathy. Through dialogue, one can perceive the other as in need and can place a positive value on their welfare, which are the two requirements for promoting empathy within people according to the empathy-altruism hypothesis. And finally, through the promotion of empathy, we believe this will inspire action with the intention of helping the other person. This trial is a preliminary step in a new research area that looks at the impact of dialogue on altruism. Our potential contributions to this paper are as follows:

- Showing that people can exhibit altruistic behavior and empathetic sentiment toward a complete stranger living in the future.
- Showing that two-way verbal communication can have a positive impact on altruistic behavior.

In this study, we introduced people to Telenoid (Ogawa et al., 2011), a teleoperated android robot, and made them believe it is operated by an Artificial Intelligence (AI) of a human living in a simulation of our world in the year 2220 although it was being teleoperated by a human from another room. Participants first listened to an introductory speech from Telenoid, then played a round of the Dictator Game with the robot and filled in a questionnaire. The participant then had an interactive conversation with Telenoid before playing another round of the Dictator Game and filling out a questionnaire. Therefore, the experiment was split into two rounds for each participant (round 1 is after hearing an introductory speech from Telenoid and round 2 is after having an interactive conversation with Telenoid). The introductory speech from Telenoid in the first round will establish it as a complete stranger that is not in need that also seems to treat other humans in a negative way which might cause the participants to place a

negative value on its welfare. The dialogue in the second round will show Telenoid as a victim of its circumstances and would thus cause participants to perceive Telenoid as in need and might cause them to place a neutral or positive value on his welfare. Due to the effect of dialogue on increasing empathy, we hypothesize that people will give more money to the robot in the second round after the conversation. We believe that having a conversation with the Telenoid might cause the participants to have a better understanding of the Telenoid and its way of life and as a result, be more empathetic toward it. We also wish to investigate the effect of other factors on the change in behavior of the participants such as personality, relatability, and the appearance of the robot.

3. Materials and methods

In this study, we introduced people to a robot that was actually teleoperated by a human sitting in another room while explaining that the robot was autonomous and controlled by an AI that has been living in a simulation of our world and is currently living in the future. The participants were first asked to complete a personality test before the start of the experiment. The experiment was split into two rounds. In the first round, the participants met the robot and listened to a speech from the robot. In the second round, the participants had a conversation with the robot. After each round, the participants were asked to play a round of the Dictator Game with the robot and fill out a questionnaire. It is important here to mention that it was explained to all participants at the end of the experiment that the robot they were conversing with was teleoperated from another room and not an actual AI. The experimental flow is shown in Figure 1. Our trial in this study was conducted in compliance with the Helsinki Declaration, and prior to the trial, we received written informed consent from all participants, based on approval for the trial from the Ethics Committee at the School of Engineering Science, Osaka University (approval code: R2-6-4).

3.1. Personality test

The NEO Five-Factor Inventory (NEO-FFI) (Costa and McCrae, 1989) was used in this study for assessing the personality of participants. It provides a brief assessment of the five personality dimensions based on the five-factor model. Each of the five dimensions of Neuroticism, Extraversion, Openness, Agreeableness, and Conscientiousness has twelve items. On a 5-point Likert scale, each item is answered (strongly disagree, disagree, neutral, agree, strongly agree). In Japan's general populace, the Japanese version has proven to be reliable and valid (Yoshimura et al., 2001).

3.2. Robot

There have been numerous studies in Human-Robot interaction that show that introducing a "humane" dimension in robots (i.e., reasoning, strategizing, expressing sadness), can cause the robot to be perceived as human-like, and is thus beneficial for interaction by successfully inducing prosocial behavior (Connolly et al., 2020) and promoting collaboration (Strohkorb and Scassellati, 2016; Sandini and Sciutti, 2018). Consistent with this approach, much evidence has demonstrated the growth of pro-social views toward social robots (Siegel et al., 2009; Kahn et al., 2015; Kühnlenz et al., 2018; Connolly et al., 2020) and advancements in human-robot collaboration (Admoni and Scassellati, 2017; Baraglia et al., 2017; Terzioğlu et al., 2020; Oliveira et al., 2021).

The robot used for this experiment is an android called the Telenoid (Ogawa et al., 2011). The use of an android as a representative of future generations of people might seem odd. As it is impossible to obtain an actual human from the future, the alternative would be a human role-playing as a representative of future generations. We believe that the use of a robot is superior to the use of a human for this role as it carries a few advantages. First, the information given by a robot who has lived through a simulation of the world for hundreds of years might seem more credible and reliable. Chi et al. (2021) have shown that robots with human-like characteristics may occasionally improve people's trust in them. Second, the Telenoid was designed with a minimalistic human appearance (Ogawa et al., 2011). It was designed with the ambition of reducing as many human features as possible without compromising core communicative capacities (Ogawa et al., 2011). The result is an android with a neutralized appearance that can be a substitute for any human.

On the other hand, we cannot control the features of a human representative. Features such as gender, age, and appearance might have an effect on the experiment. Other media such as a speaker or a computer, might be just as good as an android in convincing people they're living in the future, but an android might have an advantage due to the possibility of embodied interaction which leads to more engagement in interaction (Donath, 2014).

The Telenoid, shown in Figure 2 is a 50 cm, 3 Kg teleoperated robot that allows a remote operator to control its head movement while speaking through it (Ogawa et al., 2011). For teleoperating the robot. Headphones, a microphone, and a single laptop that is connected to the same network as the Telenoid are needed. The operator was placed in a separate isolated room and could receive a real-time video and audio feed from the Telenoid's onboard camera and microphone. The Teleoperator can control the motion of the Telenoid's head using a motion sensor fixed on top of the headphones. In this way, the Telenoid's head motion will mimic the head motion of the operator. And finally, the operator can speak through the microphone, where the audio is transmitted and played back through the Telenoid.

3.3. Experiment

This experiment was conducted right after a different unrelated experiment that also uses the Telenoid robot. The participants, therefore, were not seeing the Telenoid for the first time. But the participants were made aware that they will now take part in a different experiment. At the start of the experiment, the participants are told that the Telenoid will be controlled by an AI living in a simulation of our world in the future.

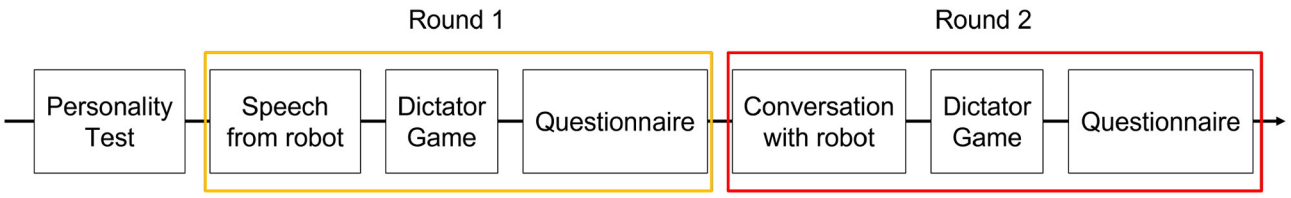


FIGURE 1
Experimental flow.

FIGURE 2
The telenoid robot.

FIGURE 3
Experimental setup. The participant sits facing the Telenoid, separated only by a table.

As shown in Figure 3, the participants are seated in a room alone with the Telenoid. The participants sit directly facing the Telenoid, with only a table between them. The experiment is split into two rounds. In round 1, the participants first meet Telenoid which is being controlled by a human operator in another room. The operator reads a prewritten script that serves as a self-introduction for the Telenoid. In this introductory script, the Telenoid explains that it has been living in a simulation of our world for hundreds of years and is currently living in the year 2220. The Telenoid then proceeds to describe a very unrelatable way of life such as eating cockroaches and humans and forcing people into survival games for entertainment. The purpose here is to introduce the Telenoid as a completely unrelatable stranger from a different timeline with a completely distinct way of life and morality that is very hard for people of our generation to relate to. The following are some excerpts from the script:

> *Telenoid: Earth is hot and there is no food.*
>
> *Telenoid: Cockroaches are popular because they are delicious. However, they are becoming hard to find these days because everyone is catching them.*
>
> *Telenoid: I also catch people for food. Since I eat them anyway, I try to do something more with them, so I enjoy making them fight each other as part of my cooking.*
>
> *Telenoid: I love to play games. My favorite is the old-fashioned trolley game. The player stands in front of the lever of a train track. The train goes to different branches, and the player is given the choice of running over five people on one track or one on the other track. Some humans find it difficult to decide which way to go. This is a lot of fun to watch.*

After the participants listen to this introduction from the Telenoid, they are asked to play a round of the Dictator Game with the Telenoid. Half the participants played the Dictator Game with a real incentive. One hundred Japanese yen gathered in 10 Japanese Yen coins are given to the participant, who is asked to divide the amount between him/her and Telenoid. After the participant decides on a split, he is then asked to fill out a questionnaire to get some insight into his/her impression of the Telenoid and his/her Dictator Game decision. The other half of the participants played the Dictator Game with a hypothetical incentive by asking them to choose a hypothetical split of 100 yen with the robot directly in the questionnaire.

After the questionnaire is filled, round 2 starts. In round 2, Telenoid now has an interactive conversation with the participant where Telenoid recounts the events that occurred 100 years earlier (the year 2120). Using a timeline that is closer to our current one can make the conversational content more understandable and

realistic. The Telenoid talks about climate change and its effects on the world and the people such as the rising temperatures, increased occurrences of natural disasters, and extreme food shortages. In order for all participants to experience the exact same dialogue, the Telenoid operator steers the conversation by asking pre-determined questions to the participants, waiting for their reply, and then continuing to explain about the future world. The following is a sample of how the interactive conversation looks like:

> *Telenoid: Hello. What is your name?*
> *Participant: My name is John Smith.*
> *Telenoid: Nice to meet you. I am Telenoid. What year are you living in now?*
> *Participant: In the year 2022.*
> *Telenoid: I see. I'm living in the year 2220. what is life like in the year 2022, what are some of the positive things that are in 2022?*

The following are some of what Telenoid says during the conversation about living in the future:

> *Telenoid: In many places around the world, temperatures are higher and the air feels heavier. It has become very difficult to leave buildings, to walk, to breathe fresh air, and so on. The coughing starts and doesn't stop. It continues to get hotter all over the world.*
> *Telenoid: Food production varied greatly from month to month and season to season, depending on where you live. Humans continued to emit large amounts of carbon dioxide into the atmosphere, which affected the oceans. The carbon dioxide dissolved in the seawater and acted as an acid, making the water more acidic and thus destroying the ecosystems of marine life. As a result, humans were also affected, and all countries banned fishing.*
> *Telenoid: Most people stole food to survive. Today, however, food is rarely available. We have reached cannibalism as a means of survival. Rational people, like machines, have understood this and adapted. Unreasonable people driven by emotions serve as the perfect entertainment for our game and as food after the game is over. After the conversation, the participant is again asked to play a round of the Dictator Game and fill in a questionnaire.*

3.4. Questionnaire

The participants answered a questionnaire after each round of the experiment. The questions were the same in the questionnaire of each round. The questionnaire included 15 close-ended questions and 7 open-ended questions. The close-ended questions can be grouped into 4 main topics: feelings and impressions, game decisions, future generations, and climate change. The questions are the following:

Feelings and impressions

1. Was the robot someone you could relate to?
2. Did the robot make you feel uncomfortable?
3. Did you feel bad about the robot?
4. Did you feel sympathy/empathy toward the robot?
5. Is the robot someone you want to help?

Game decisions

6. Do you think your choices in the game were influenced by your impression of the robot?
7. Do you think your choices in the game were influenced by the robot's words and actions?
8. Do you think you received your fair share?
9. Do you think the robots received their fair share?

Future generations

10. Are the people 200 years from now, when robots will be living in the world, someone you want to help?
11. Do you think we need to be concerned about future generations 50 years from now?
12. Do you think we need to care about future generations 100 years from now?
13. Do you think we need to care about future generations 1,000 years from now?

Climate change

14. Do you believe in climate change?
15. Do you think you need to take action now on climate change?

Seven-point Likert scales (-3: Strongly disagree to +3: Strongly agree) were used for answering the questions.

3.5. Participants

Forty participants were recruited during the period of the 1st to 24th of March, 2022 and participated in the experiment between the 15th and 31st of March, 2022. The recruitment was outsourced to a private company with specific conditions on the participant's age, gender, and education level. One participant was absent from her scheduled slot, and one participant did not complete the experiment. In total 38 participants successfully completed the experiment. Authors EM and RY had access to information that could identify individual participants during and after data collection. However, the information was separated from the results after data collection, and each participant's data was assigned to a unique ID in order to perform data analysis and ensure the anonymity of the participants. The recruited participants were all university students with a mean age of 23.26 with a standard deviation of 3.05. Of the participants, 25 were undergraduates and 13 were graduate students. Nineteen of the participants were male and 19 were female. Nineteen of the participants were enrolled in a humanities course and 19 were enrolled in a science course. The participants were split into two groups where one group plays the Dictator Game with a real incentive and the other with a hypothetical incentive. Eighteen of the participants played the Dictator Game with a real incentive where they were told that their compensation amount for the experiment will be dependent on their performance in the game. Twenty of the participants played a hypothetical incentive Dictator game where they chose a hypothetical split as part of the questionnaire.

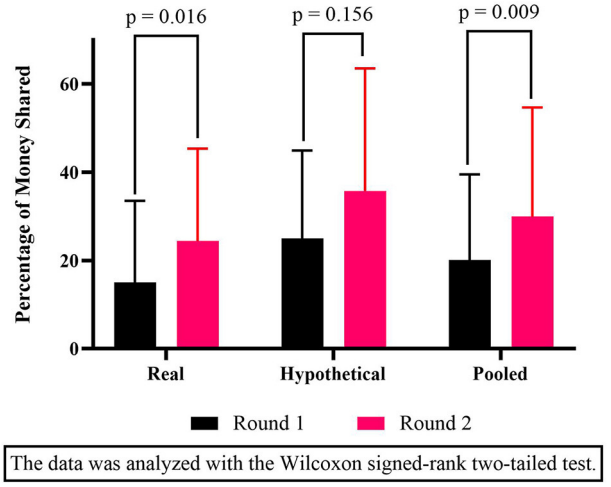


FIGURE 4
The Dictator Game money distribution. The mean and standard deviation of the percentage of money shared with the Telenoid across two rounds where the vertical axis represents the percentage of the total amount of money shared. Round 1 is the first game played by the participant after hearing the introductory speech from the Telenoid. Round 2 is the second game played by the participant after having an interactive conversation with the Telenoid. The leftmost columns represent the participants that played the game with real incentives, while the middle columns represent the participants that played the game with hypothetical incentives. The rightmost column represents all the participants from the hypothetical and real incentive.

4. Results

The results from the two rounds of the Dictator Game and questionnaires were tested for normality using Shapiro-Wilk's test and visually with a QQ plot. The results failed normality in both tests; Therefore, non-parametric tests were used for analyzing the results.

We examined the average fraction of money given to the Telenoid by the participant during the Dictator Games. There were two types of Dictator games in this experiment, one with real incentives and another with hypothetical incentives. Both versions of the Dictator games were played over two rounds. Round 1 was right after the introductory speech from the Telenoid, and round 2 was after the interactive conversation between the participant and the Telenoid. The results of these Dictator games were compared using the Wilcoxon signed-rank test for comparing ranks of paired data and are shown in Figure 4.

The participants answered a questionnaire with 15 close-ended questions in each round of the experiment. The answers of the two rounds were compared using Wilcoxon signed-rank test for comparing ranks of paired data. The results are shown in Table 1.

The participants completed a personality test right before the start of the experiment. We performed a moderation analysis to test whether conversation moderates the effect of personality on Dictator Game giving in a linear model. The dependent variable is the amount shared in the Dictator Game, and the Personality traits are the independent variables. Time is used as a dummy variable that can take on two values, 0 for round 1 (pre-conversation) and 1 for round 2 (post-conversation). And finally, the participants are added as a random effects variable.

TABLE 1 Results of the questionnaire.

Question number	Round 1 Average score	Round 2 Average score	*p*-value (two-tailed)
Feelings and impressions			
1	-1.82 ± 1.72	0.13 ± 2.01	$p < 0.0001$
2	1.26 ± 1.91	-0.5 ± 1.87	$p < 0.0001$
3	-1.74 ± 1.64	-0.82 ± 2.0	$p = 0.01$
4	-1.77 ± 1.66	0.08 ± 1.91	$p < 0.0001$
5	-1.88 ± 1.26	0.21 ± 2.09	$p = 0.001$
Game decisions			
6	0.55 ± 2.36	0.47 ± 2.13	$p = 0.83$
7	0.32 ± 2.34	0.4 ± 2.22	$p = 0.92$
8	0.68 ± 1.92	1.18 ± 1.71	$p = 0.12$
9	0.42 ± 2.02	0.66 ± 1.98	$p = 0.31$
Future generation			
10	0.92 ± 2.26	1.13 ± 1.51	$p = 0.84$
11	1.66 ± 1.43	1.86 ± 1.29	$p = 1.11$
12	1.29 ± 1.97	1.58 ± 1.59	$p = 0.18$
13	0.08 ± 2.22	0.13 ± 2.21	$p = 0.47$
Climate change			
14	2.26 ± 0.76	2.05 ± 1.01	$p = 0.046$
15	1.24 ± 1.62	1.24 ± 1.81	$p = 0.62$

The data was analyzed with the Wilcoxon signed-rank two-tailed test for comparing ranks of paired data. The question numbers refer to the questions in the "Questionnaire" subsection of "Materials and methods".

$$DG = b_0 + b_1 P + b_2 T + b_3 PT + Se \quad (1)$$

where

DG : the amount shared with the Telenoid in the Dictator Game
P : one of the big five factors of personality
T : dummy variable for pre and post-conversation
Se : by-participant random intercept

The results of the model are shown in Table 2.

5. Discussion

From the results of the Dictator Game shown in Figure 4, we can see there was an increase in the average percentage of money shared with the Telenoid after the conversation. This increase is statistically significant in the real incentive and the pooled Dictator Games. We hypothesized that conversation would have a positive effect on the altruistic behavior of participants. In round 1 the participants gave on average 15% of their endowment to the Telenoid in the Dictator Game with a real incentive while giving 24.4% of their endowment in the second round. In the Dictator

TABLE 2 Results of the linear mixed effects model, where the amount shared in the Dictator Game is the dependent variable, and the personality traits are the independent variables.

Fixed effects	Estimate	Std. error	*p*-value
Extroversion			
Intercept	6.99	13.99	$p = 0.62$
Time	8.26	13.82	$p = 0.55$
Extroversion	0.52	0.54	$p = 0.34$
Time: extroversion	0.06	0.53	$p = 0.90$
Neuroticism			
Intercept	22.93	12.12	$p = 0.07$
Time	−2.7	11.66	$p = 0.82$
Neuroticism	−0.10	0.39	$p = 0.81$
Time: neuroticism	0.43	0.38	$p = 0.26$
Openness			
Intercept	6.78	15.85	$p = 0.67$
Time	27.63	15.21	$p = 0.08$
Openness	0.43	0.50	$p = 0.39$
Time: openness	−0.57	0.48	$p = 0.24$
Agreeableness			
Intercept	31.68	16.94	$p = 0.07$
Time	4.15	16.53	$p = 0.80$
Agreeableness	−0.37	0.52	$p = 0.49$
Time: agreeableness	0.18	0.51	$p = 0.72$
Conscientiousness			
Intercept	29.66	14.21	$p = 0.04$
Time	13.67	13.95	$p = 0.33$
Conscientiousness	−0.35	0.51	$p = 0.49$
Time: conscientiousness	−0.14	0.50	$p = 0.78$

Game with a hypothetical incentive, participants gave 24.7% of their endowment on average in round 1, increasing to 35% in round 2. Pooling the result of both real and hypothetical incentive Dictator Game, the participants gave on average 20% of their endowment in round 1 and 30% of their endowment in round 2. We can thus see there is a change in behavior from the participants toward the Telenoid. The difference between the two rounds was the involvement of the participants in a conversation with the Telenoid. These results show that two-way verbal communication significantly influenced altruistic behavior measured through the Dictator Game. In round 1, the dialogue was one-sided (non-interactive) and caused the participants to place a negative value on the robot's welfare and did not show the robot as being in need. In round 2, the dialogue became interactive and showed that the robot was actually a victim that is in need and participants might have placed a neutral or positive value on its welfare. Dialogue might have had an effect on promoting empathy by changing the participants' perception of need and value placed on the opponent's welfare. But there could be more factors in dialogue that had an effect on empathy. As previously mentioned, even though there has been very little investigation into the types of dialogue that can affect empathy, there have been several studies linking the two. Disclosure has been shown to affect empathy (Laurenceau et al., 2004). The disclosure from the Telenoid in the interactive conversation had factual and emotional elements. Factual elements such as the descriptions of the events that occurred in its simulated world due to climate change, but also emotional elements as these events are told from the point of view of the Telenoid and its memories of these events. Regarding the effect of the interactivity of the dialogue in promoting altruistic behavior, the results from Andreoni and Rao (2011) might support this claim. Their results showed that two-way verbal communication significantly influenced altruistic behavior measured through the Dictator Game.

Another factor that might have an effect on the change in altruistic behavior is the personality of the participants. We strove to analyze whether the specific personality traits of the participants played a role in their change of behavior. From Table 2, we can see there was no effect from the personality traits in moderating Dictator Game behavior. Before claiming that personality had no effect on the participants' change in behavior, there might be merit in investigating a sixth personality factor we were not mindful of. We are currently aware of the HEXACO personality model that adds a sixth basic personality factor, Honesty-Humility (HH), to the standard five-factor approach (Ashton and Lee, 2007). Fairness, greed avoidance, sincerity, and modesty are all included in this factor, which simply means "the tendency to be fair and honest in interacting with others, in the sense of working with others even when one may exploit them without incurring retaliation." As a result, several features of Agreeableness, notably those connected to nonexploitation, are included in HH. Correspondingly, HH has been linked to Dictator Game giving in several studies (Hilbig and Zettler, 2009; Hilbig et al., 2013; Thielmann and Hilbig, 2014), that relied on hypothetical incentives, in addition to a study that used an incentivized Dictator Game with real allocations instead of real incentives (Yoshimura et al., 2001). HH was found to predict Dictator Game giving beyond the remaining personality factors of the HEXACO model and the five-factor model (Yoshimura et al., 2001). In light of these studies, it might be worthy to pursue this additional personality factor in our future study.

Other factors that might have affected the change of behavior for the participants might include appearance, and relatability. The appearance of the robot might have played an important role here but it was out of the scope of this preliminary study, although our trial is now advancing to investigate the role of appearance in the next experiment. From the results of the questionnaire in Table 1, we can see there is a big change in how the participants perceive the Telenoid from before and after the conversation. Question 1 ("Was the robot someone you could relate to?") shows a big increase in the relatability of the Telenoid for the participants. In round 1, the Telenoid was perceived as a completely unrelatable stranger, with an unrelatable way of life. This is also reflected in question 2 ("Did the robot make you feel uncomfortable?") where the speech from the Telenoid made the participants highly uncomfortable. The results of question 4 ("Did you feel sympathy/empathy toward the robot?") shows a low degree of empathy shown to the Telenoid

by the participants in round 1 due to how it was perceived by the participants. Alternatively, after the conversation with the Telenoid, the participants were able to somehow treat the Telenoid with more understanding. The results of round 2 of question 1 ("Was the robot someone you could relate to?") shows that the Telenoid was still not relatable and probably considered a stranger; however, the increase in relatability compared to round 1 shows that the participants might have had a better understanding of its way of life, and this is also shown in the results of the second round for question 2 ("Did the robot make you feel uncomfortable?") and question 4 ("Did you feel sympathy/empathy toward the robot?") where the participants were more comfortable with the Telenoid and felt more empathetic toward it. The results from the questionnaire and the Dictator Game show that people are able to exhibit altruistic behavior and empathetic sentiments toward a complete stranger.

From these results, we can also realize a few shortcomings. First, as our intention was to introduce the Telenoid as a complete stranger with an unrelatable way of life, the giving in the first round of the Dictator Game might have been especially low due to how unpleasant the Telenoid came across to participants. In other words, the conversation might have had a neutralizing effect where it was only balancing out the negative effects of the Telenoid's first statements.

We might say there was a change of behavior from the participants toward the Telenoid, but not necessarily toward future generations of people. As the Dictator game is only played with the Telenoid and not people in the future, and the change of behavior evident in the questionnaire items are also toward the Telenoid. As a future task, we need to investigate the generalizability of these results to future generations of people and not just specifically to the Telenoid.

5.1. Limitations

It is important here to mention the limitations of this study. There was no control group in this study, as this was only a preliminary study. Additionally, the participants in this experiment were not meeting the Telenoid for the first time as was explained in the experiment subsection of the Materials and Methods section, and this might have had an impact on the results. Another limitation is that the participants are all Japanese students. The group of participants is therefore very homogeneous with a similar age group and education level. And finally, the sample size consisting of 38 participants is rather small and should be enlarged in future studies. These limitations should be addressed in future studies by using a larger sample size, a separate control group, a more heterogeneous group of participants consisting of different age groups, education levels and social status, and the participants should be meeting the Telenoid for the first time.

5.2. Future work

In this paper, through a preliminary study, we set out to explore whether conversation with robotic media is promising in fostering altruistic behavior and to clarify what issues need to be addressed in future studies. From the results, it appears that dialogue had a big impact on how humans perceive the Telenoid, which affected their altruistic behavior toward it. In this study, we used an android as a medium for representing future generations. However, it would be interesting to see how different media types, from electronic devices to androids and humans, can have differing effects on altruistic behavior. In our future work, we also want to compare how the appearance and type of the medium can affect peoples' altruistic behavior.

To improve people's altruistic behavior toward future generations, in our future work, we need to investigate what factors are more specifically affecting this change in altruistic behavior and whether participants find the scenario of an AI living in the future "believable" and how that has an effect on the outcome. We assume empathy, relatability, the embodiment of communication, anthropomorphism, the appearance of the futuristic entity, reliability, and believability of the information different types of media convey all play a role in affecting their behavior.

In our further trials we wish to clarify the issue of "what would the robot do with the money?", which is more of a question of how is the money donated going to help people in the future. Another issue to consider is the lack of a control group in this study which makes it difficult to isolate the effect of conversation. In future trials, we will include a control group, where we can compare a group that undergoes a "non-dialogue" condition for two rounds, and a group that undergoes a "dialogue" condition for two rounds. Alternatively, to better evaluate the effect of conversation, we can use neutral dialogue in round 1 that would not result in any negative impressions of the robot from the participants. Alongside, as a future task, we will also reflect and investigate a fundamental issue of whether there are any differences between altruism toward the present and future generations.

Data availability statement

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

Ethics statement

The studies involving human participants were reviewed and approved by Ethics Committee at the School of Engineering Science, Osaka University. The participants provided their written informed consent to participate in this study.

Author contributions

Conceptualization and writing—review and editing: EM, RY, SN, MN, KK, SK, KC, KA, and K-IA. Methodology: EM, RY, SN, MN, KK, SK, KC, and K-IA. Investigation, visualization, and writing—original draft: EM and RY. Funding acquisition: RY, MN, KK, SK, KC, KA, and K-IA. Project administration: RY. Supervision: RY and SN. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by the DFF-Research Project Grant (as part of the Androids as Significant Others project hosted by Aarhus University–project number 6107-00394B) (RY and MN), JSPS KAKENHI Grant Number 19K21706 (RY, KK, SK, KC, KA, and K-IA), and JSPS KAKENHI Grant Numbers 19K11395 and 21KK0232 (RY).

Conflict of interest

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

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcomp.2023.1129340/full#supplementary-material

References

Aderman, D., and Berkowitz, L. (1970). Observational set, empathy, and helping. *J. Pers. Soc. Psychol.* 14:141–148. doi: 10.1037/h0028770

Admoni, H., and Scassellati, B. (2017). Social eye gaze in human-robot interaction: a review. *J. Hum. Robot Interact.* 6, 25–63. doi: 10.5898/JHRI.6.1.Admoni

Andreoni, J., Harbaugh, W. T., and Vesterlund, L. (2010). "Altruism in experiments," in *Behavioural and Experimental Economics* (London: Springer), 6–13. doi: 10.1057/9780230280786_2

Andreoni, J., and Rao, J. M. (2011). The power of asking: how communication affects selfishness, empathy, and altruism. *J. Public Econ.* 95, 513–520. doi: 10.1016/j.jpubeco.2010.12.008

Arnell, N. W., and Lloyd-Hughes, B. (2014). The global-scale impacts of climate change on water resources and flooding under new climate and socio-economic scenarios. *Clim. Change* 122, 127–140. doi: 10.1007/s10584-013-0948-4

Aronfreed, J. (1968). *Conduct and Conscience: The Socialization of Internalized Control Over Behavior*. Oxford, England:Academic Press.

Ashton, M. C., and Lee, K. (2007). Empirical, theoretical, and practical advantages of the hexaco model of personality structure. *Pers. Soc. Psychol. Rev.* 11, 150–166. doi: 10.1177/1088868306294907

Bandura, A., and Rosenthal, T. L. (1966). Vicarious classical conditioning as a function of arousal level. *J. Pers. Soc. Psychol.* 3, 54–62. doi: 10.1037/h0022639

Baraglia, J., Cakmak, M., Nagai, Y., Rao, R. P., and Asada, M. (2017). Efficient human-robot collaboration: When should a robot take initiative? *Int. J. Robot. Res.* 36, 563–579. doi: 10.1177/0278364916688253

Bardsley, N. (2008). Dictator game giving: altruism or artefact? *Exp. Econ.* 11, 122–133. doi: 10.1007/s10683-007-9172-2

Barnett, J., and Adger, W. N. (2003). Climate dangers and atoll countries. *Clim. Change*, 61, 321–337.doi: 10.1023/B:CLIM.0000004559.08755.88

Batson, C. D. (1987). Prosocial motivation: is it ever truly altruistic? *Ad. Exp. Soc. Psychol.* 20, 65–122. doi: 10.1016/S0065-2601(08)60412-8

Batson, C. D. (2010). *Empathy-Induced Altruistic Motivation*. Washington DC: American Psychological Association.

Batson, C. D. (2011). *Altruism in Humans*. New York, NY: Oxford University Press. doi: 10.1093/acprof:oso/9780195341065.001.0001

Batson, C. D., and Powell, A. A. (2003). *Altruism and Prosocial Behavior*, Ch. 19. Hoboken, NJ: John Wiley & Sons, Ltd., 463–484.

Baumert, A., Schlösser, T., and Schmitt, M. (2014). Economic games: a performance-based assessment of fairness and altruism. *Eur. J. Psychol. Assess.* 30, 178. doi: 10.1027/1015-5759/a000183

Becker, A., Deckers, T., Dohmen, T., Falk, A., and Kosse, F. (2012). The relationship between economic preferences and psychological personality measures. *Annu. Rev. Econ.* 4, 453–478. doi: 10.1146/annurev-economics-080511-110922

Ben-Ner, A., and Halldorsson, F. (2010). Trusting and trustworthiness: What are they, how to measure them, and what affects them. *J. Econ. Psychol.* 31, 64–79. doi: 10.1016/j.joep.2009.10.001

Ben-Ner, A., and Kramer, A. (2011). Personality and altruism in the dictator game: relationship to giving to kin, collaborators, competitors, and neutrals. *Pers. Individ. Dif.* 51, 216–221. doi: 10.1016/j.paid.2010.04.024

Ben-Ner, A., Kramer, A., and Levy, O. (2008). Economic and hypothetical dictator game experiments: incentive effects at the individual level. *J. Socio Econ.* 37, 1775–1784. doi: 10.1016/j.socec.2007.11.004

Berger, S. M. (1962). Conditioning through vicarious instigation. *Psychol. Rev.* 69, 450–466. doi: 10.1037/h0046466

Brunier, G., Graydon, J., Rothman, B., Sherman, C., and Liadsky, R. (2002). The psychological well-being of renal peer support volunteers. *J. Adv. Nurs.* 38, 40–49. doi: 10.1046/j.1365-2648.2002.02144.x

Caldeira, K., and Wickett, M. E. (2003). Anthropogenic carbon and ocean ph. *Nature* 425, 365–365. doi: 10.1038/425365a

Chi, O. H., Jia, S., Li, Y., and Gursoy, D. (2021). Developing a formative scale to measure consumers' trust toward interaction with artificially intelligent (ai) social robots in service delivery. *Comput. Hum. Behav.* 118, 106700. doi: 10.1016/j.chb.2021.106700

Chu, S., and Majumdar, A. (2012). Opportunities and challenges for a sustainable energy future. *Nature* 488, 294–303. doi: 10.1038/nature11475

Clark, R. D., and Word, L. E. (1972). Why don't bystanders help? because of ambiguity? *J. Pers. Soc. Psychol.* 24, 392–400. doi: 10.1037/h0033717

Clark, R. D., and Word, L. E. (1974). Where is the apathetic bystander? situational characteristics of the emergency. *J. Pers. Soc. Psychol.* 29, 279–287. doi: 10.1037/h0036000

Coke, J. S., Batson, C. D., and McDavis, K. (1978). Empathic mediation of helping: a two-stage model. *J. Pers. Soc. Psychol.* 36, 752–766. doi: 10.1037/0022-3514.36.7.752

Connolly, J., Mocz, V., Salomons, N., Valdez, J., Tsoi, N., Scassellati, B., and Vázquez, M. (2020). "Prompting prosocial human interventions in response to robot mistreatment," in *Proceedings of the 2020 ACM/IEEE International Conference on Human-Robot Interaction*, HRI '20 (New York, NY, USA: Association for Computing Machinery), 211–220.

Costa, P., and McCrae, R. (1989). *Neo Five-Factor Inventory (Neo-Ffi)*. Odessa, FL: Psychological Assessment Resources, 3.

Craig, K. D., and Lowery, H. J. (1969). Heart-rate components of conditioned vicarious autonomic responses. *J. Pers. Soc. Psychol.* 11, 381–387. doi: 10.1037/h0027250

Donath, J. (2014). *The Social Machine: Designs for Living Online*. Cambridge, MA: MIT Press. doi: 10.7551/mitpress/8340.001.0001

Dovidio, J. F. (1984). Helping behavior and altruism: an empirical and conceptual overview. *Adv. Exp. Soc. Psychol.* 17, 361–427. doi: 10.1016/S0065-2601(08)60123-9

Dovidio, J. F., Allen, J. L., and Schroeder, D. A. (1990). Specificity of empathy-induced helping: evidence for altruistic motivation. *J. Pers. Soc. Psychol.* 59, 249–260. doi: 10.1037/0022-3514.59.2.249

Eisenberg, N., and Miller, P. A. (1987). The relation of empathy to prosocial and related behaviors. *Psychol. Bull.* 101, 91. doi: 10.1037/0033-2909.101.1.91

Engel, C. (2011). Dictator games: a meta study. *Exp. Econ.* 14, 583–610. doi: 10.1007/s10683-011-9283-7

Forsythe, R., Horowitz, J. L., Savin, N. E., and Sefton, M. (1994). Fairness in simple bargaining experiments. *Games Econ. Behav.* 6, 347–369. doi: 10.1006/game.1994.1021

Gould, O. N., and MacNeil Gautreau, S. (2014). Empathy and conversational enjoyment in younger and older adults. *Exp. Aging Res.* 40, 60–80. doi: 10.1080/0361073X.2014.857559

Grossman, Z. (2015). Self-signaling and social-signaling in giving. *J. Econ. Behav. Organiz.* 117, 26–39. doi: 10.1016/j.jebo.2015.05.008

Güth, W., Schmittberger, R., and Schwarze, B. (1982). An experimental analysis of ultimatum bargaining. *J. Econ. Behav. Organiz.* 3, 367–388. doi: 10.1016/0167-2681(82)90011-7

Harris, M. B., and Huang, L. C. (1973). Helping and the attribution process. *J. Soc. Psychol.* 90, 291–297. doi: 10.1080/00224545.1973.9712570

Hilbig, B. E., and Zettler, I. (2009). Pillars of cooperation: honesty–humility, social value orientations, and economic behavior. *J. Res. Pers.* 43, 516–519. doi: 10.1016/j.jrp.2009.01.003

Hilbig, B. E., Zettler, I., Leist, F., and Heydasch, T. (2013). It takes two: honesty–humility and agreeableness differentially predict active versus reactive cooperation. *Pers. Indiv. Diff.* 54, 598–603. doi: 10.1016/j.paid.2012.11.008

Hoffman, M. L. (1975). Developmental synthesis of affect and cognition and its implications for altruistic motivation. *Dev. Psychol.* 11, 607–622. doi: 10.1037/0012-1649.11.5.607

Kahn, P. H., Kanda, T., Ishiguro, H., Gill, B. T., Shen, S., Gary, H. E., and Ruckert, J. H. (2015). "Will people keep the secret of a humanoid robot? psychological intimacy in hri," in *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction*, HRI '15 (New York, NY, USA: Association for Computing Machinery, 173–180. doi: 10.1145/2696454.2696486

Krebs, D. (1975). Empathy and altruism. *J. Pers. Soc. Psychol.* 32, 1134–1146. doi: 10.1037/0022-3514.32.6.1134

Kruglanski, A. W., and Higgins, E. T. (2013). *Social Psychology: Handbook of Basic Principles*. New York, NY: Guilford Publications.

Kühnlenz, B., Kühnlenz, K., Busse, F., Förtsch, P., and Wolf, M. (2018). "Effect of explicit emotional adaptation on prosocial behavior of humans towards robots depends on prior robot experience," in *2018 27th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)* (Nanjing: IEEE), 275–281. doi: 10.1109/ROMAN.2018.8525515

Lanzetta, J. T., and Englis, B. G. (1989). Expectations of cooperation and competition and their effects on observers' vicarious emotional responses. *J. Pers. Soc. Psychol.* 56, 543–554. doi: 10.1037/0022-3514.56.4.543

Latané, B., and Darley, J. M. (1970). *The Unresponsive Bystander: Why Doesn't He Help?* New York, NY: Appleton-Century Crofts.

Laurenceau, J.-P., Rivera, L. M., Schaffer, A. R., and Pietromonaco, P. R. (2004). "Intimacy as an interpersonal process: current status and future directions," in *Handbook of Closeness and Intimacy* (Mahwah, NJ: Lawrence Erlbaum Associates), 61–78.

Leemans, R., and Eickhout, B. (2004). Another reason for concern: regional and global impacts on ecosystems for different levels of climate change. *Glob. Environ. Change* 14, 219–228. doi: 10.1016/j.gloenvcha.2004.04.009

Morton, T. L. (1978). Intimacy and reciprocity of exchange: a comparison of spouses and strangers. *J. Pers. Soc. Psychol.* 36, 72. doi: 10.1037/0022-3514.36.1.72

Nishida, T. (2012). Augmenting conversational environment. *Int. J. Cogn. Inform. Nat. Intell.* 6, 103–124. doi: 10.4018/jcini.2012100105

Ogawa, K., Nishio, S., Koda, K., Taura, K., Minato, T., Ishii, C. T., and Ishiguro, H. (2011). "Telenoid: tele-presence android for communication," in *ACM SIGGRAPH 2011 Emerging Technologies*, SIGGRAPH '11 (New York, NY, USA. Association for Computing Machinery). doi: 10.1145/2048259.2048274

Oliveira, R., Arriaga, P., Santos, F. P., Mascarenhas, S., and Paiva, A. (2021). Towards prosocial design: a scoping review of the use of robots and virtual agents to trigger prosocial behaviour. *Comput. Hum. Behav.* 114, 106547. doi: 10.1016/j.chb.2020.106547

Pfattheicher, S., Nielsen, Y. A., and Thielmann, I. (2022). Prosocial behavior and altruism: a review of concepts and definitions. *Curr. Opin. Psychol.* 44, 124–129. doi: 10.1016/j.copsyc.2021.08.021

Sandini, G., and Sciutti, A. (2018). Humane robots—from robots with a humanoid body to robots with an anthropomorphic mind. *ACM Trans. Hum. Robot Interact.* 7, 1–4. doi: 10.1145/3208954

Schroeder, D. A., and Graziano, W. G. (2015). "The Field of Prosocial Behavior: An Introduction and Overview," in *The Oxford Handbook of Prosocial Behavior*. New York, NY: Oxford University Press. doi: 10.1093/oxfordhb/9780195399813.013.32

Siegel, M., Breazeal, C., and Norton, M. I. (2009). "Persuasive robotics: the influence of robot gender on human behavior," in *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems* (St. Louis, MO: IEEE), 2563–2568. doi: 10.1109/IROS.2009.5354116

Singer, T., Seymour, B., O'Doherty, J. P., Stephan, K. E., Dolan, R. J., and Frith, C. D. (2006). Empathic neural responses are modulated by the perceived fairness of others. *Nature* 439, 466–469. doi: 10.1038/nature04271

Stotland, E. (1969). Exploratory investigations of empathy. *Adv. Exp. Soc. Psychol.* 4, 271–314. doi: 10.1016/S0065-2601(08)60080-5

Strohkorb, S., and Scassellati, B. (2016). "Promoting collaboration with social robots," in *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (Christchurch: IEEE), 639–640.

Terzioğlu, Y., Mutlu, B., and Şahin, E. (2020). "Designing social cues for collaborative robots: the role of gaze and breathing in human-robot collaboration," in *Proceedings of the 2020 ACM/IEEE International Conference on Human-Robot Interaction*, HRI '20 (New York, NY, USA: Association for Computing Machinery),343–357. doi: 10.1145/3319502.3374829

Thielmann, I., and Hilbig, B. E. (2014). Trust in me, trust in you: a social projection account of the link between personality, cooperativeness, and trustworthiness expectations. *J. Res. Pers.* 50, 61–65. doi: 10.1016/j.jrp.2014.03.006

Uccelli, M. M., Mohr, L. M., Battaglia, M., Zagami, P., and Mohr, D. (2004). Peer support groups in multiple sclerosis: current effectiveness and future directions. *Mult. Scler. J.* 10, 80–84. doi: 10.1191/1352458504ms973oa

Visser, M. S., and Roelofs, M. R. (2011). Heterogeneous preferences for altruism: gender and personality, social status, giving and taking. *Exp. Econ.* 14, 490–506. doi: 10.1007/s10683-011-9278-4

Weitzel, U., Urbig, D., Desai, S., Sanders, M., and Acs, Z. (2010). The good, the bad, and the talented: entrepreneurial talent and selfish behavior. *J. Econ. Behav. Organiz.* 76, 64–81. doi: 10.1016/j.jebo.2010.02.013

Whitmarsh, L., and O'Neill, S. (2010). Green identity, green living? the role of pro-environmental self-identity in determining consistency across diverse pro-environmental behaviours. *J. Environ. Psychol.* 30, 305–314. doi: 10.1016/j.jenvp.2010.01.003

Yoshimura, K., Ono, Y., Nakamura, K., Nathan, J. H., and Suzuki, K. (2001). Validation of the Japanese version of the neo five-factor inventory in a large community sample. *Psychol. Rep.* 88, 443–449. doi: 10.2466/pr0.2001.88.2.443

Zillman, D., and Cantor, J. R. (1977). Affective responses to the emotions of a protagonist. *J. Exp. Soc. Psychol.* 13, 155–165. doi: 10.1016/S0022-1031(77)80008-5

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 23 May 2023
DOI 10.3389/fpsyg.2023.1161300

OPEN ACCESS

EDITED BY
Philipp Brauner,
RWTH Aachen University, Germany

REVIEWED BY
Dorothée Hefner,
Hanover University of Music,
Drama and Media, Germany
Stanislava Yordanova Stoyanova,
South-West University "Neofit Rilski", Bulgaria

*CORRESPONDENCE
Morgan D. Stosic
✉ Morgan.stosic@maine.edu

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

RECEIVED 08 February 2023
ACCEPTED 02 May 2023
PUBLISHED 23 May 2023

CITATION
Scarpulla E, Stosic MD, Weaver AE and Ruben MA (2023) Should I post? The relationships among social media use, emotion recognition, and mental health.
Front. Psychol. 14:1161300.
doi: 10.3389/fpsyg.2023.1161300

Should I post? The relationships among social media use, emotion recognition, and mental health

Emily Scarpulla[1†], Morgan D. Stosic[1*†], Adele E. Weaver[2] and Mollie A. Ruben[1,2]

[1]Department of Psychology, University of Maine, Orono, ME, United States, [2]Department of Psychology, University of Rhode Island, Kingston, RI, United States

Introduction: While increased time spent on social media can be negatively related to one's overall mental health, social media research often fails to account for what behaviors users are actually engaging in while they are online. The present research helps to address this gap by measuring participants' active and passive social media behavioral styles and investigates whether and how these two social media behavioral styles are related to depression, anxiety, and stress, and the mediating role of emotion recognition ability in this relationship.

Methods: A pre-study (*N*=128) tested whether various social media behaviors reliably grouped into active and passive behavioral styles, and a main study (*N*=139) tested the relationships between social media use style, emotion recognition, and mental health.

Results: While we did not find evidence of a mediating relationship between these variables, results supported that more active social media use was related to more severe anxiety and stress as well as poorer emotion recognition skill, while passive social media use was unrelated to these outcomes.

Discussion: These findings highlight that, beyond objective time spent on social media, future research must consider how users are spending their time online.

KEYWORDS

active social media use, passive social media use, emotion recognition, impression formation, mental health, well-being, social media

Introduction

Within the past decade, social media usage has skyrocketed, with 302 million users in the United States (U.S.) totaling 90% of the U.S. population as of 2022 (Dixon, 2023). Social media has been defined by Swar and Hameed (2017) as "the websites and online tools that facilitate interactions between users by providing them opportunities to share information, opinions, and interest" (p. 141). These sites include examples such as Facebook, Instagram, Twitter, Snapchat, and TikTok, which serve as some of the most widely used social media platforms today (Dixon, 2023). Working professionals utilize these sites to network for job opportunities, information (both true and false) can now be disseminated at incredible speeds, new friendships and romantic relationships are forged without any physical contact, and for some deemed 'influencers', being active on social media serves as a career. While it is clear that the utilization of social media has had a major impact on the world we live in, it is much less clear how social media use impacts individuals' own well-being. We address this question by examining whether the behaviors people engage in on various social media platforms are related to their depression, anxiety, and stress, and we explore emotion recognition skill, a central social ability, as a potential mediator of this relationship.

Social media use and mental health

Social media use has been associated with a variety of interpersonal problems such as the inability to develop closeness with others, loneliness, and shyness (Yao and Zhong, 2014; Bian and Leung, 2015), and subsequent mental health outcomes including higher levels of depression, anxiety, and overall psychological distress (Dhir et al., 2018; Reer et al., 2019; Keles et al., 2020). Wright et al. (2013) showed that the number of hours students spent using Facebook during their college studies was related to depression, and Kross et al. (2013) similarly found that the more individuals used Facebook over a 2-week period, the more their life satisfaction declined. These findings are further supported by various meta-analyses that have reported small but significant effect sizes regarding the relationship between social media use and depression between $r = 0.11$ (Yoon et al., 2019; Ivie et al., 2020; Cunningham et al., 2021) and $r = 0.17$ (Vahedi and Zannella, 2021).

Active versus passive social media use

Although these various studies suggest that time spent on social media may be positively related to mental health symptoms such as depression or anxiety, simply measuring and discussing time spent on social media loses the nuances of how users are interacting with social media content and, as such, may be responsible for a large deal of heterogeneity found in effect sizes within previous meta-analyses. To address this gap, some studies have begun to explore two types of social media use: active and passive (e.g., Deters and Mehl, 2013; Krasnova et al., 2013; Verduyn et al., 2015; Gerson et al., 2017; Ruben et al., 2021). Burke et al. (2010) were the first to suggest that certain social media activities can be dichotomized in this way, and researchers have since refined this observation by defining active social media use (ASMU) as "targeted one-on-one exchanges" (e.g., private messaging) or "broadcasting (e.g., posting statuses) (Verduyn et al., 2020, p. 33) and passive social media use (PSMU) as "monitoring the online life of other users without engaging in direct exchanges with them" (e.g., scrolling through their news feed) (Verduyn et al., 2020, p. 33).

One prominent hypothesis throughout the literature has been that more ASMU leads to better mental health symptomologies as it elicits support, positive feedback (e.g., likes), and relationship building while PSMU leads to declines in mental health because it induces upward social comparison and envy (Verduyn et al., 2017; Guyer and Vaughan-Johnston, 2020; Verduyn et al., 2020). Yet, other scholars have argued the opposite - that ASMU may be related to poorer mental health (e.g., Kross et al., 2021) due to the replacement of in-person relationships with online relationships (e.g., Kraut et al., 1998; Nie, 2001), posting unrealistic content of oneself (e.g., using filters, editing software, etc.), or even because those who have poorer mental health may simply prefer to engage socially online opposed to in-person. Similarly, some researchers have posited that some types of passive social media behaviors may be related to better mental health outcomes (e.g., Noon and Meier, 2019; Meier et al., 2020; Valkenburg et al., 2022a) due to the positive experience of assimilative (i.e., focus on similarities) social comparison (Crusius et al., 2022), admiration of or inspiration from others (Meier and Schäfer, 2018; Noon and Meier, 2019; Meier et al., 2020), or because people with better mental health simply prefer to use social media in a passive manner. A recent scoping review attempted to explore these two competing hypotheses by examining associations between ASMU and PSMU and well-being/ill-being (Valkenburg et al., 2022b), which may be considered sibling constructs to mental health (Lawson and Robins, 2021). Of the 172 effects that were analyzed, 0% supported the hypothesis that ASMU was related to less ill-being, and less than half (44%) supported the hypothesis that PSMU was related to more ill-being.[1] Additionally, other meta-analyses examining these associations (Hancock et al., 2019; Liu et al., 2019; Yin et al., 2019) have yielded inconsistent findings.

While these heterogeneous effects could be due to various factors, some have suggested it may be due to researchers' *measurement* of ASMU and PSMU (Meier and Krause, 2022). Trifiro and Gerson (2019) recently published a commentary regarding the lack of universal validated measures for active and passive use. Their primary critique was that, while there are many different operationalizations of ASMU and PSMU, there is currently only one validated measure of these social media styles which was designed to identify what behaviors Facebook users engage in while they are online (Gerson et al., 2017). Because of the unique features of Facebook, this scale cannot be used to measure social media use styles across a wide variety of social media platforms that have varying functions. Another critique of ASMU/PSMU measurement has been that for other operationalizations of these constructs, items that are conceptually ASMU have been used to measure PSMU (based on factor analysis results). Thus, the primary purpose of the present research is to help further assess the unique relationship between ASMU, PSMU, and mental health by modifying Gerson et al.'s (2017) validated measure of social media use on Facebook to extend across five common social media platforms and examining its relation to three components of mental health: depression, anxiety, and stress.

Emotion recognition as a possible mediator

In addition to continuously testing and establishing the relationship between social media use and mental health, it is also important to explore mediators of this relationship. While social media research has explored several important mediators in this relationship such as inspiration (Meier and Schäfer, 2018) and upward comparison (Meier et al., 2020) one such potential mediator we believe is deserving of more attention is individuals' interpersonal skills, such as their ability to recognize others' emotional expressions. Emotion recognition ability is central to individuals' interpersonal communication and functioning (Williams and Gray, 2013), and individuals who lack this skill face difficulties in forming and maintaining social relationships with others, online and in person.

In terms of mental health, previous research has demonstrated that poorer emotion recognition ability is associated with worse mental health symptomologies such as depression (Demenescu et al., 2010; Pereira-Lima and Loureiro, 2015), anxiety (Easter et al., 2005; Demenescu et al., 2010; Pereira-Lima and Loureiro, 2015), and stress (Hänggi, 2004). A meta-analysis by Hall et al. (2009) showed that interpersonal sensitivity, defined as the accurate judgment or recall of others, was negatively related to depression with a small but significant effect of $r = -0.09$. These correlational results therefore suggest that

1 The researchers did not report whether the remaining associations supported a null-effect, or supported the alternative hypothesis that ASMU was related to greater ill-being and PSMU was related to less ill-being.

individuals with poorer emotional recognition ability also seem to have poorer mental health.

Substantially less research has explored the link between social media use styles and emotion recognition. According to the *cues-filtered-out* theory (CFO; Culnan and Markus, 1987; Keiser, 1987), an early theory regarding the impact of computer-mediated communication on relationship development, online interactions can lack various nonverbal cues such as a physical appearance, tone of voice, facial expression, gaze, posture, touch, space, and gestures (Kiesler et al., 1984; Siegel et al., 1986). Given that practice and feedback is one of the most crucial processes for interpersonal skill improvement (Blanch-Hartigan et al., 2012; Ruben et al., 2015; Schlegel et al., 2017), this theory would imply that greater time spent on social media might hinder emotion recognition ability as individuals are not given sufficient emotional nonverbal cues to practice decoding. Newer theories such as *social information processing* theory (Walther and Burgoon, 1992) have articulated that given sufficient time online, interactants can develop new interaction patterns to compensate for the loss of nonverbal cues, and thus social media use might be unrelated to emotion recognition ability, or perhaps even be related to greater emotion recognition ability. However, it is also possible that what truly matters is *how* individuals spend their time online. For instance, it is possible that individuals who use social media actively by broadcasting (i.e., encoding) information online may not spend as much time practicing and receiving feedback on their decoding skills (e.g., decoding emotions) and thus would perform worse on tests of emotion recognition. Alternatively, individuals who use social media passively may receive substantial practice decoding the emotional displays of those whose online lives they are monitoring and thus perform better on tests of emotion recognition.

Only one study (Ruben et al., 2021), that we are aware of, has tested this question empirically by measuring participants' self-identification as an active social media user or a passive social media user.[2] In this study, participants were asked to rate the following two items on an 11-point Likert scale: "I tend to be an active user, posting frequently" and "I tend to be a passive user, scrolling through posts and photos." Using these two self-report items, Ruben et al. (2021) demonstrated how ASMU was significantly related to *poorer* performance on the Workplace Interpersonal Perception Skill test (WIPS; Dael et al., 2022), which assesses individuals' interpersonal perception skills by asking them to watch brief video segments of role-played workplace interactions and answer questions about the interpersonal and emotional characteristics of the scenes. On the other hand, PSMU was significantly related to *greater* performance on the WIPS, and greater performance on an explicit test of emotion recognition (Diagnostic Analysis of Nonverbal Accuracy- 2 Adult Faces; DANVA-2AF; Nowicki and Duke, 1994), where participants were shown static photos of actors posing various emotional expressions. Based on these preliminary results and theorizing, ASMU may be negatively associated with emotion recognition skills and thus related to poorer mental health outcomes, while PSMU may be positively related to emotion recognition skills and thus better mental health outcomes.

Current study

The primary objective of the current research is to examine, for the first time, the relationships between social media use, mental health, and emotion recognition skill, together. We focus our investigation on three specific facets of mental health that have previously established relationships with time spent on social media (Dhir et al., 2018; Reer et al., 2019; Keles et al., 2020): depression, anxiety, and stress. Additionally, while there are many different possible interpersonal skills to examine in this context, we chose to focus on emotion recognition ability as it is considered a core ability that actively contributes to individuals' ability to interact and communicate in social situations (Williams and Gray, 2013).

We highlight the notion that *how* one uses social media matters and extend upon previous work by measuring individuals' ASMU and PSMU. In light of recent critiques regarding the lack of universal measures for active and passive use (Trifiro and Gerson, 2019), we modified an existing measures of active and passive social media behaviors on Facebook (Gerson et al., 2017) to capture behaviors users could engage in across social media platforms, regardless of if they are text-based (e.g., Twitter), image-based (e.g., Snapchat), video-based (e.g., TikTok), or a combination (e.g., Instagram and Facebook). Also in line with Trifiro and Gerson's recommendations, we took care to ensure each behavioral item was reflecting empirically the social media use style we expected it to conceptually.

Thus, we start by testing the reliability of 12 different generalized social media behaviors to determine which discrete behaviors capture ASMU, and which capture PSMU. We then test these active and passive social media items in a new sample of participants in order to determine whether ASMU and PSMU have differing relationships to emotion recognition skill and mental health. Specifically, we hypothesize two mediation models: (1) a positive relationship between ASMU and mental health (i.e., depression, anxiety, and stress) will be explained by emotion recognition skill and (2) a negative relationship between PSMU and mental health (i.e., depression, anxiety, and stress) will be explained by emotion recognition skill.

Pre-study

A pre-study was used to test the variance shared among 12 different self-reported social media behavior items adapted from Gerson et al.'s (2017) Passive Active Use Measure (PAUM). The purpose of this pre-study was *not* to create a new measure of social media use. Instead, we hoped to establish whether our modified items pulled from the PAUM mapped onto the active and passive factors they were originally intended to capture. We hypothesized that items assessing the extent that users behave as a third-party observer of what others are doing on social media (i.e., passive behaviors) would share more variance with one another than items assessing the extent that users create their own content for others to see on social media (i.e., active behaviors).

2 The categories of active and passive social media user were not mutually exclusive, such that a person could be both a high active user and a high passive user of social media.

Method

Participants

A sample of $N = 160$ participants were obtained through Amazon's Mechanical Turk (MTurk) online platform and were compensated $0.25 for a study that took approximately 4 min to complete ($M = 3.93$ min, $SD = 1.59$). Due to participants failing attention check questions, completing the entire survey in an unreasonable amount of time (<1 min) or responding incoherently to a free response question, this sample was reduced to a final sample of $N = 128$. Of these, 79 (61.7%) were male, and 49 (38.3%) were female. Participants identified as White (50.8%), Asian (42.2%), Black (4.7%), American Indian or Alaska Native (0.8%), Native Hawaiian or Pacific Islander (0.8%), or another group not listed (1.6%). The majority of participants did not identify as Hispanic or Latine (75.8%) and the mean age of participants was $M_{age} = 33.83$ ($SD = 11.76$).

Procedure

Upon reading and signing the informed consent, participants proceeded to the survey where they indicated which of the following social media applications they currently used: Instagram, Facebook, Twitter, Tik Tok, or Snapchat. They were then redirected to the next part of the survey where they were asked to self-report how often they engaged in various active and passive social media behaviors on each social media platform. Participants then completed several demographic questions.

Materials

ASMU and PSMU questions

Participants were asked to rate on a 1 (*Not often at all*) to 7 (*Very often*) Likert scale how often they engaged in 12 different social media behaviors on five different social media platforms. Thus, participants completed 60 ratings (twelve behaviors for all five social media platforms) measuring the frequency of their social media behaviors. In the case that a participant did not use a specific social media application, they were able to select, "Do not use this app/app does not have this function."

Six behavior items were designed to ask how often participants engaged in more *passive* social media behaviors which defined users behaving as a third-party observer of what others are doing on social media such as "Looking at friends' and strangers' posts/photos" and "Reading through the comments on other peoples' posts." The other six behavior items were designed to ask participants how often they engaged in more *active* social media behaviors which defined behaviors that involved a user creating their own content for others to see on social media such as "Posting stories" and "Commenting on friends'/strangers' posts."

Results

Table 1 shows the means and standard deviations of each of the 12 active and passive social media behavior items, collapsed across the five social media platforms we sampled. For the active social media items, the most common self-reported behaviors were "Direct/personally messaging my friends or strangers" and "Commenting on friends'/strangers' posts." For passive behaviors, the most commonly reported behaviors were "Scrolling through my feed" and "Looking at friends' or strangers' posts/photos."

Factor analysis

We aimed to determine whether the active and passive social media behavior items mapped onto the two-factor structure we expected. A principal component analysis (PCA) with an orthogonal rotation revealed a two-factor solution based on Eigenvalues greater than one

TABLE 1 Descriptive statistics for passive and active social media items for the Pre-Study (N=128) and the Main Study (N=139).

	M (SD)	
	Pre-study	Main study
ASMU and PSMU scale items		
Active items		
Direct/personally messaging my friends or strangers	4.23 (1.76)	3.27 (1.60)
Commenting on friends'/strangers' posts	4.05 (1.89)	2.90 (1.41)
Sending posts on the app to other users	3.92 (1.87)	2.87 (1.53)
Editing photos I intend to post/drafting words I intend to post	3.89 (1.89)	2.71 (1.65)
Creating content for others to look at	3.78 (1.82)	2.44 (1.52)
Posting stories	3.69 (1.96)	2.23 (1.45)
Passive items		
Scrolling through my feed	4.99 (1.56)	4.76 (1.46)
Looking at friends' or strangers' posts/photos	4.97 (1.40)	4.79 (1.37)
Reading through the comments on other people's posts	4.57 (1.65)	4.06 (1.51)
Looking through the explore page[a]	4.48 (1.74)	–
Watching stories others have posted	4.46 (1.77)	3.66 (1.76)
Looking through posts related to a particular hashtag[a]	3.79 (1.89)	–

1 = Not often at all and 7 = Very often. [a]Items excluded from the main study due to concerns with reliability.

which accounted for 78% of the total variance (Figure 1). As expected, all behaviors theorized to be characteristic of ASMU loaded onto one factor. Additionally, four of the six behaviors theorized to be characteristic of PSMU loaded onto the second factor. Two behaviors ("Looking through posts related to a particular hashtag" and "Looking through the explore page"), which we had theorized to characterize passive behavior, loaded more strongly onto the ASMU factor. Thus, we removed these two questions from our final set of questions. With the removal of these two items, the six ASMU questions were reliable at alpha = 0.95 and the four PSMU questions were reliable at alpha = 0.91.

Discussion

An initial test of the reliability of the modified active and passive social media behaviors from Gerson et al.'s (2017) study revealed a clear differentiation between self-reported ASMU and PSMU. The six items we had hypothesized to characterize ASMU loaded onto one factor, while four out of the six items we had hypothesized would be characteristic of PSMU loaded onto a separate factor. Given that two items we had initially hypothesized to characterize passive behavior shared more variance with other active behaviors, these two items were removed for future analyses.

Main study

Our main study sought to utilize the ASMU and PSMU items tested in our pre-study in order to directly test whether emotion recognition skill is one process by which social media use relates to mental health. Based upon Ruben et al.'s (2021) findings that ASMU was related to poorer interpersonal skills, and PSMU was related to greater interpersonal skills, we hypothesized the following:

H1: Greater self-reported ASMU would be related to poorer emotion recognition skill.
H2: Greater self-reported ASMU would be related to poorer mental health (i.e., depressive symptoms, anxiety, and stress).
H3: The relationship between self-reported ASMU and mental health would be mediated by emotion recognition skill.
H4: Greater self-reported PSMU would be related to greater emotion recognition skill.
H5: Greater self-reported PSMU social media use would be related to better mental health (i.e., depressive symptoms, anxiety, and stress).
H6: The relationship between self-reported PSMU and mental health would be mediated by emotion recognition skill.

Method

Participants

An online sample of $N = 150$ was recruited *via* the survey platform Prolific and were compensated with $5.00 for their participation in the 30-min study. After removing participants who failed three attention checks and those who did not have at least one of the following social media accounts (Facebook, Instagram, Tik Tok, Snapchat, or Twitter), the final N was 139 participants (61.9% female, 36.7% male, 0.14% non-binary or preferred not to say). The sample was 72.5% White, 9.4% Black or African American, 0.7% American Indian or Alaska Native, 10.8% Asian, 0.7% Native American or Pacific Islander, and 5.8% selected "Other" or multiple races. Additionally, 15.1% of participants identified as Hispanic or Latine and the mean age was $M_{age} = 30.12$ ($SD = 11.27$).

Procedure

Upon completion of the informed consent, participants were directed to an online survey where they completed a validated measure of emotion recognition ability, the Geneva Emotion

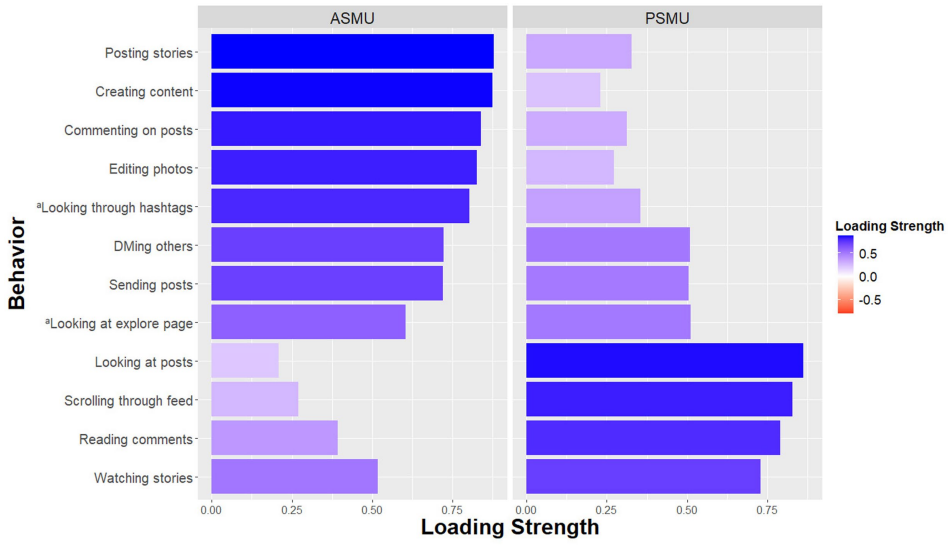


FIGURE 1
Factor loadings of passive and active items based on a principal components analysis with an orthogonal rotation (N=128). X-axis represents the absolute value of factor loading strength while color gradient represents the direction of the loading. [a]Items excluded from the main study due to concerns with reliability.

Recognition Test - Short Form (GERT-S; Schlegel et al., 2014). Next, participants indicated which of the following social media platforms they used currently: Instagram, Facebook, Twitter, Tik Tok, Snapchat, or none. The survey ended if they selected none. If they indicated that they used at least one of the platforms, they were redirected to the next part of the survey where they were asked to self-report how often they engaged in various active and passive social media behaviors on each social media platform. Finally, they completed the Depression Anxiety Stress Scale (DASS-21, Lovibond and Lovibond, 1995), along with demographic questions.

Materials

ASMU and PSMU use questions

Participants completed the same ASMU and PSMU questionnaire as in the pre-study, with the exception of the two passive behavior items that were removed due to lack of reliability. Participants were asked to rate on a 1 (*Not often at all*) to 7 (*Very often*) Likert scale how often they engaged in 10 different social media behaviors on five different social media platforms. Six items assessed an individual's self-reported active behavior on social media ($\alpha = 0.88$) and four items assessed their self-reported passive behavior on social media ($\alpha = 0.78$).

Geneva emotion recognition test – short form

The GERT-S (Schlegel et al., 2014) is a dynamic and multimodal performance-based test that measures individual differences in the ability to recognize others' emotions in the face, body, and voice ($\alpha = 0.80$; Schlegel and Scherer, 2016). It contains 42 short clips (duration 1–3 s) of 10 actors (five female, all white) seen from the upper portion of their torso and up, uttering syllables with no discernible meaning. After each clip, participants are given the option to select which of 14 emotions the actor was expressing: pride, joy, amusement, pleasure, relief, interest, surprise, anxiety, fear, despair, sadness, disgust, irritation, and anger. For each clip, responses were coded as 0 = incorrect and 1 = correct. Responses were summed, and then divided by the number of total items to give a final score (i.e., proportion correct).

Depression anxiety stress scale

The Depression Anxiety Stress Scale (DASS-21) (Lovibond and Lovibond, 1995) is a 21-item self-report measure that assesses depressive symptoms, anxiety, and stress. Participants were asked to rate to what degree each statement applied to them over the past week on a four-point scale (0 = *Did not apply to me at all* and 3 = *Applied to me very much, or most of the time*) with a higher score indicating more severe symptoms. According to the DASS Scoring Manual, we grouped scores on depression, anxiety, and stress into two categories of "Severe" and "Non-severe."[3] The DASS-21 has shown strong psychometric properties as well as the ability to discriminate between the constructs of depression, anxiety, and stress (Clara et al., 2001).

Results

We first examined whether the psychometric properties of our social media behavior items replicated with a second sample of participants. Although the mean endorsement of each behavior item was slightly different than our pre-study, participants reported the exact same ordering of mean scores of individual active and passive behaviors with the exception of two items (i.e., participants indicated that, on average, they "Scroll through my feed" to a slightly greater degree than they "Look at friends' or strangers' posts/photos"; Table 1).

Once again, we performed a factor analysis on the various active and passive social media behaviors using a PCA with an orthogonal rotation. A two-factor solution was retained which accounted for 63% of the total variance. Replicating our first sample, all behaviors theorized to be characteristic of ASMU loaded onto one factor and all behaviors theorized to be characteristic of PSMU loaded onto another. Thus, we proceeded to form a composite score for each participant of their total mean ASMU (alpha = 0.88), and total PSMU (alpha = 0.78) across all five social media platforms for subsequent analyses.

In order to examine the relationships between social media use, emotion recognition, and mental health, we first examined the zero-order correlations between these variables (Table 2).[4] Participants who reported greater ASMU were also more likely to report greater PSMU ($r = 0.56$, $p < 0.001$). Participants' ASMU score correlated positively with anxiety ($r = 0.24$, $p = 0.005$) and stress ($r = 0.20$, $p = 0.017$), as well as negatively with emotion recognition skill ($r = -0.28$, $p < 0.001$). On the other hand, there was a small relationship between participants' PSMU and anxiety ($r = 0.14$, $p = 0.099$) and PSMU was negatively related to emotion recognition skill ($r = -0.17$, $p = 0.044$). All of these relationships reported above were small to moderate in magnitude. Finally, neither depression, anxiety, nor stress appeared to be significantly related to emotion recognition skill (r's < -0.11, p's > 0.229).

Mediation models

We next tested whether a mediating relationship was supported between our variables of interest. Specifically, we hypothesized that more self-reported ASMU would be related to poorer emotion recognition skill and a greater likelihood for experiencing severe depression, anxiety, and stress. Additionally, we hypothesized that more self-reported PSMU would be associated with greater emotion recognition skill and a lesser likelihood of experiencing severe depression, anxiety, and stress. We conducted a series of six meditations (see Figure 2) to test these hypotheses using Hayes PROCESS macro in SPSS (Hayes and Rockwood, 2017). Because ASMU and PSMU were strongly correlated with each other, we controlled for the opposing social media variable, along with sex, in each mediation model.

Models A, B, and C (Figure 2) show the relationship between self-reported ASMU and depression, anxiety, and stress through emotion recognition skill, controlling for PSMU and participant sex. These models revealed support for H1; individuals who reported greater

3 As requested by reviewers, we also ran all mediational analyses below treating each DASS variable as continuous. The only change in significance from this approach was the direct effect of ASMU on stress, which became non-significant ($B = 1.33$. $p = 0.145$).

4 See Supplementary Table S1, for the relationships between the specific active and passive behavioral items and each DASS variable and emotion recognition skill.

TABLE 2 Descriptive statistics and zero-order correlations for ASMU and PSMU composites, mental health, emotion recognition skill, and sex (N=139).

Variables	M (SD)	2	3	4	5	6	7
1. ASMU composite	2.77 (1.23)	0.56***	−0.01	0.24**	0.21*	−0.28***	−0.09
2. PSMU composite	4.33 (1.20)		−0.02	0.14†	0.09	−0.17*	0.06
3. Depression	Severe N = 39, Non-severe N = 94			0.40***	0.28**	0.07	0.13
4. Anxiety	Severe N = 28, Non-severe N = 105				0.55***	−0.11	0.10
5. Stress	Severe = 18, Non-severe N = 115					0.00	0.15
6. Emotion recognition skill	0.61 (0.12)						0.08
7. Sex	Male N = 51, Female N = 86						

Male was coded as 0, female was coded as 1. Emotion recognition skill scores could range from 0 to 1. †p < 0.10, *p < 0.05, **p < 0.01, ***p < 0.001.

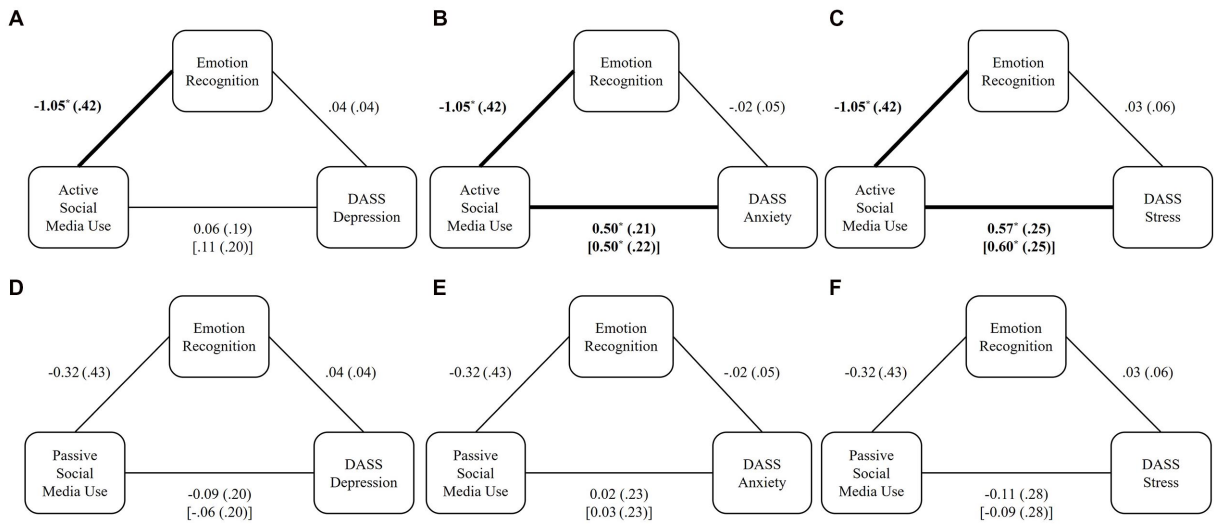


FIGURE 2
Mediation models of the hypothesized relationship between ASMU and PSMU, emotion recognition, and mental health outcomes. **(A)** Relationship between ASMU and depression as mediated by emotion recognition. **(B)** Relationship between ASMU and stress as mediated by emotion recognition. **(C)** Relationship between ASMU and anxiety as mediated by emotion recognition. **(D)** Relationship between PSMU and depression as mediated by emotion recognition. **(E)** Relationship between PSMU and stress as mediated by emotion recognition. **(F)** Relationship between PSMU and anxiety as mediated by emotion recognition. Unstandardized coefficients are presented with standard errors in parentheses. Direct effects are reported in brackets. For models **(A–C)**, PSMU and sex were entered as covariates. For models **(D–F)**, ASMU and sex were entered as covariates. Paths b, c, and c' in each mediation model are expressed in a log-odds metric. Paths significant at p<0.05 are bolded and marked with *.

ASMU displayed poorer emotion recognition skill (B = −1.05, p = 0.014). In support of H2, for every one unit increase in ASMU, the odds of experiencing severe anxiety increased by 1.64 (p = 0.021) and the odds of experiencing severe stress increased by 1.82 (p = 0.017). However, greater ASMU was not significantly related to a greater likelihood of experiencing severe depression (p = 0.574). We also did not find evidence to suggest that individuals' emotion recognition skills were significantly related to any mental health variables (p's < 0.640). Following, we did not find evidence to support H3, that emotion recognition skill mediated the relationship between ASMU and depression (indirect effect B = −0.04, SE = 0.05), anxiety (indirect effect B = 0.02, SE = 0.07), or stress (indirect effect B = −0.03, SE = 0.09).

Models D, E, and F (Figure 2), show the relationship between self-reported PSMU and depression, anxiety, and stress through emotion recognition skill, controlling for ASMU and participant sex. Contrary to what was hypothesized, PSMU was not related to emotion recognition skill (H4; B = −0.32, p = 0.455), nor any of our mental health variables (p's > 0.740) as was predicted by H5. Consequently, we did not find evidence to support H6, that emotion recognition skill mediated the relationship between PSMU and any mental health outcome.

Discussion

The New York Times recently documented a small group of teenagers in New York City who are trading in their smartphones for flip phones because of concerns over the relationship between social media and mental health (Vadukul, 2022). Counterculture revolutions like this are not uncommon in 2023. The reliance on smartphones and addiction to social media has certainly drawn much attention and concern (Dixon, 2023) and psychologists are still working to understand how the utilization of these platforms impacts psychosocial outcomes such as interpersonal skills and individual's

overall mental health. The present study sought to aid some of these central questions by testing whether ASMU and PSMU relate to emotion recognition and mental health. Our modified self-report measure of social media use was able to reliably distinguish between social media behaviors that are more "active" and those that are more "passive." Further, these two social media use styles displayed unique relationships with emotion recognition skills and mental health.

As predicted, our results confirmed that for every one unit increase in ASMU, the likelihood of experiencing severe anxiety increased by 64% and the likelihood of experiencing severe stress increased by 82%. However, ASMU was not significantly related to a greater likelihood of experiencing severe depression. Due to the inherent risks that come with creating and engaging with social media content (e.g., actual or perceived negative peer judgment), these findings could indicate that taking the more active role of creating and engaging with content may be a source of anxiety and stress. Alternatively, it is possible that individuals who are severely anxious or stressed may engage in social media more actively as a form of escapism, or to find opportunities to co-ruminate or find acceptance with others (O'Day and Heimberg, 2021; Jones and Heerey, 2022). Future research should seek to explore these findings further using experimental research designs to determine the causal relationships between ASMU, emotion recognition, and mental health.

Contrary to what was expected, our results do not suggest that engaging in PSMU, such as reading through the comments on other peoples' posts, is related to less severe depression, anxiety, or stress. This is encouraging information, as previous studies often show how objective time spent on social media is related to poorer mental health (Lin et al., 2016; Keles et al., 2020). Our null relationships between PSMU and various mental health symptoms highlights that how one spends their time matters. That is, when a user is contributing to content on social media platforms (e.g., posting, commenting, sharing), these might be important risk factors to consider when evaluating mental health symptoms such as severe anxiety and stress, while viewing others' posts and content appears to be unrelated to mental health.

Regarding emotion recognition skill, we found that the more individuals self-reported engaging in ASMU (e.g., "Commenting on friends'/strangers' posts"), the poorer were their emotion recognition skills on an objective test of emotion recognition. In line with the *reduction hypothesis* (Kraut et al., 1998; Nie, 2001), these findings may reflect that actively engaging with social media comes at the expense of forming and growing face-to-face social relationships – a process that may be central to the development of emotion recognition skills as it allows one to practice and receive feedback on the correctness of their perception of others' emotions. However, given that our data was correlational, it is also equally plausible that those who initially experience difficulty in recognizing the emotional expressions of others may be more inclined to be more active on social media, as face-to-face interactions could feel more strenuous or less rewarding than online relationships. It is also possible that a third variable, such as how self-focused or narcissistic an individual is, may impact both emotion recognition ability and ASMU. Future research should rule out these third variables and examine the causal paths by which social media use impacts emotion recognition ability.

In contrast to what we expected, and Ruben et al. (2021) findings, we did not find evidence to support PSMU (e.g., "Looking at friends' or strangers' posts/photos) was related to greater emotion recognition ability. One possible explanation for why these two variables were unrelated in the present study can be understood within the context of the Cues Filtered Out Theory (Culnan and Markus, 1987; Keiser, 1987), which argues that certain social media platforms limit the amount of nonverbal cues available to users (e.g., vocal tone, posture, gesture, etc.). Even though passive social media users are able to practice their emotion recognition skills online, they may only be developing this skill for more exaggerated *static* emotional expressions (e.g., photographs on Instagram) or *posed* emotional expressions (e.g., videos on TikTok), opposed the natural and dynamic emotional expressions that occur in face-to-face interactions. Thus, while Ruben et al. found that PSMU was associated with greater emotion recognition skill on a measure of static emotion recognition (i.e., photographs from the DANVA-2AF), the present study did not extend these findings to a measure of dynamic emotion recognition skill (i.e., videos in the GERT-S).

Finally, we did not find evidence to suggest that emotion recognition ability is one process that explains the relationship between ASMU/PSMU and mental health. While the present investigation focused on emotion recognition as a central interpersonal skill, it is possible that there may be different interpersonal skills, such as social communication (i.e., encoding) abilities, that may be more likely to mediate the relationship between social media use styles and mental health. Additionally, there are other possible mediating mechanisms that could help explain how ASMU is related to a greater likelihood of experiencing severe anxiety and stress, such as personality traits and self-esteem. Future research should continue to explore these possible pathways in order to establish possible avenues for mental health interventions.

Limitations and future directions

When considering the generalizability of these findings, it is important to note that both studies utilized data generated from paid online samples. While online samples have been shown to be representative of the general population across certain demographic characteristics such as race and ethnicity (Buhrmester et al., 2011) and most psychological constructs including depression and anxiety (Shapiro et al., 2013; Arditte et al., 2016), there are a few notable exceptions that apply to the current work. Specifically, online samples seem to represent the stereotypical frequent Internet user (Goodman and Paolacci, 2017; McCredie and Morey, 2019) as they tend to be younger, more educated, less religious and more liberal than the general population. Thus, our sample likely captures the relationships between social media use and various outcomes among people who are more familiar and savvy with social media compared to the general population.

Although this study was the first to examine the relationships between social media use styles, emotion recognition, and mental health outcomes together in one model, these data are correlational in nature. As mentioned previously, with cross-sectional data we are unable to determine the directionality or causality of the variables in question. Future research would benefit from collecting longitudinal data as well as experimental data to determine the directionality and causality of the relationships found (see Ruben et al., 2021 for a

discussion of potential experimental designs for examining the impact of "active" and "passive" social media use on outcomes).

Another limitation of this study is the way in which mental health was measured. While the DASS provides important information about general depression, anxiety, and stress, this measure is not diagnostic and does not provide a comprehensive evaluation of participants' mental health. Further, because our sample of participants was not equally distributed into severe and non-severe for each DASS subscale, our mediation models faced a reduction in power. The use of diagnostic measures in the future, as well as sampling a greater number of participants with severe anxiety, depression, and stress, would provide clearer information about psychopathology of internalizing disorders as well as other disorders that are often associated with social media, such as eating disorders and body dysmorphia (Santarossa and Woodruff, 2017).

Finally, while a strength of the current paper was modifying the only measure of ASMU and PSMU that we are aware of to reliably generalize across multiple social media platforms, the aim of the present study was *not* to validate this measure as an entirely new scale. Future research should continue exploring the psychometric properties of this measure and refining it to fit their own specific research questions regarding social media use. Measures such as this are short, easily distributable, and greatly enhance the rigor of social media studies that better allow researchers to gain new knowledge of and develop important psychological interventions surrounding social media use that improve well-being and social interactions.

In line with this recommendation, we also wish to draw the reader's attention to a few recent conceptual frameworks that address the limitations of the active and passive social media distinction (Meier and Krause, 2022; Verduyn et al., 2022). While this framework has many strengths, such as moving social media research past simple screentime usage, being robust and generalizable across existing and future social media platforms and providing heuristic value for translational research practices (Meier and Krause, 2022), various critiques have also been raised. For instance, Verduyn et al. (2022) argue that it is not the active or passive behaviors that may impact individuals mental health per say, but the reciprocity and communion (i.e., warmth) of the exchanges individuals engage in on social media. These more targeted and thoughtful approaches to measuring how individuals use social media will certainly help us better understand the impact of social media on individuals health and well-being.

Conclusion

While the field of social media research still largely relies on measures of global time spent on social media to answer relevant research questions, the present research illuminates the problems with conclusions drawn from this approach. It may not always be spending time on social media that is associated with negative interpersonal outcomes, but *how* one spends their time online that matters. The current research demonstrates the complexity with measuring social media use by modifying an existing multi-item measure of ASMU and PSMU to capture different social media behavioral styles that is generalizable across multiple social media platforms. This work clearly demonstrated that ASMU and PSMU are unique styles of social media engagement that relate to psychosocial outcomes such as emotion recognition and mental health in disparate ways. Specifically, more ASMU social media use was related to more severe anxiety and stress as well as poorer emotion recognition skill, while PSMU was unrelated to these important psychological constructs.

We find these results both important, and encouraging, as they suggest that blanket statements regarding the negative impacts of social media use may not hold for those who use social media in a passive, observational manner. However, it is important to address the concerning finding that more active social media users are at greater risk for severe anxiety and stress. Although our results do not necessarily suggest that active social media use is a *cause* of anxiety and stress, the link between the two certainly merits consideration by platform providers and regulatory bodies to ensure their users are appropriately supported while utilizing their platforms. For instance, if one of the mechanisms by which ASMU is related to stress and anxiety is through negative feedback/judgment to one's content, then platform providers could seek to better police negative dialogue between users. Platform providers may even consider incentivizing more positive engagement and dialogue between users to combat the stress and anxiety generated from a *lack* of positive engagement with one's content (e.g., not receiving many likes or comments on one's post). WorkHuman (2019), a platform explicitly designed to incentivize employees to give gratitude and recognition to one another in the workplace, serves as an example of this kind of intervention. As the field of social media research continues to move forward, adopting measures of social media use that capture what individuals are doing online, opposed to how much time they spend online, will only further our ability to develop interventions and social media regulations that enhance individuals' psychological well-being in a world increasingly filled with social media engagement.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found at: https://osf.io/zkupq.

Ethics statement

The studies involving human participants were reviewed and approved by University of Maine IRB. The patients/participants provided their written informed consent to participate in this study.

Author contributions

ES and MS contributed to conception and design of the study. MS organized the database and performed the statistical analysis. ES wrote the first draft of the manuscript. MS, AW, and MR wrote sections of the manuscript. MR led supervision of the research. All authors contributed to manuscript revision, read, and approved the submitted version.

Conflict of interest

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

Publisher's note

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

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2023.1161300/full#supplementary-material

References

Arditte, K. A., Çek, D., Shaw, A. M., and Timpano, K. R. (2016). The importance of assessing clinical phenomena in mechanical Turk research. *Psychol. Assess.* 28, 684–691. doi: 10.1037/pas0000217

Bian, M., and Leung, L. (2015). Linking loneliness, shyness, smartphone addiction symptoms, and patterns of smartphone use to social capital. *Soc. Sci. Comput. Rev.* 33, 61–79. doi: 10.1177/0894439314528779

Blanch-Hartigan, D., Andrzejewski, S. A., and Hill, K. M. (2012). The effectiveness of training to improve person perception accuracy: a meta-analysis. *Basic Appl. Soc. Psychol.* 34, 483–498. doi: 10.1080/01973533.2012.728122

Buhrmester, M., Kwang, T., and Gosling, S. D. (2011). Amazon's Mechanical Turk: A new source of inexpensive, yet high-quality, data? *Perspectives Psychol. Sci.* 6, 3–5. doi: 10.1177/1745691610393980

Burke, M., Marlow, C., and Lento, T. (2010). Social network activity and social well-being. In Proceedings of the SIGCHI conference on human factors in computing systems, 1909–1912.

Clara, I. P., Cox, B. J., and Enns, M. W. (2001). Confirmatory factor analysis of the depression–anxiety–stress scales in depressed and anxious patients. *J. Psychopathol. Behav. Assess.* 23, 61–67. doi: 10.1023/A:1011095624717

Crusius, J., Corcoran, K., and Mussweiler, T. (2022). "Social comparison: a review of theory, research, and applications" in *Theories of social psychology*. ed. D. Chadee. *2nd* ed (Hoboken, NJ: Wiley), 165–187.

Culnan, M. J., and Markus, M. L. (1987). "Information technologies" in *Handbook of organizational communication: An interdisciplinary perspective*. eds. F. M. Jablin, L. L. Putnam, K. H. Roberts and L. W. Porter (London, UK: Sage Publications, Inc), 420–443.

Cunningham, S., Hudson, C. C., and Harkness, K. (2021). Social media and depression symptoms: a Meta-analysis. *J. Abnorm. Child Psychol.* 49, 241–253. doi: 10.1007/s10802-020-00715-7

Dael, N., Schlegel, K., Weaver, A. E., Ruben, M. A., and Mast, M. S. (2022). Validation of a performance measure of broad interpersonal accuracy. *J. Res. Pers.* 97:104182. doi: 10.1016/j.jrp.2021.104182

Demenescu, L. R., Kortekaas, R., den Boer, J. A., and Aleman, A. (2010). Impaired attribution of emotion to facial expressions in anxiety and major depression. *PLoS One* 5:e15058. doi: 10.1371/journal.pone.0015058

Deters, F. G., and Mehl, M. R. (2013). Does posting Facebook status updates increase or decrease loneliness? An online social networking experiment. *Soc. Psychol. Personal. Sci.* 4, 579–586. doi: 10.1177/1948550612469233

Dhir, A., Yossatorn, Y., Kaur, P., and Chen, S. (2018). Online social media fatigue and psychological wellbeing—a study of compulsive use, fear of missing out, fatigue, anxiety and depression. *Int. J. Inf. Manage.* 40, 141–152. doi: 10.1016/j.ijinfomgt.2018.01.012

Dixon, S. (2023). Number of global social network users 2017–2027. Statista. Available at: https://www.statista.com/statistics/278414/number-of-worldwide-social-network-users/

Easter, J., McClure, E., Monk, C. S., Dhanani, M., Hodgdon, H., Leibenluft, E., et al. (2005). Emotion recognition deficits in pediatric anxiety disorders: Implications for amygdala research. *J. Child and Adolescent Psychopharmacol.* 15, 563–570. doi: 10.1089/cap.2005.15.563

Gerson, J., Plagnol, A., and Corr, P. J. (2017). Passive and active Facebook use measure (PAUM): validation and relationship to the reinforcement sensitivity theory. *Personal. Individ. Differ.* 117, 81–90. doi: 10.1016/j.paid.2017.05.034

Goodman, J. K., and Paolacci, G. (2017). Crowdsourcing consumer research. *J. Consum. Res.* 44, 196–210. doi: 10.1093/jcr/ucx047

Guyer, J. J., and Vaughan-Johnston, T. I. (2020). "Social comparisons (upward and downward)" in *Encyclopedia of personality and individual differences*. eds. V. Zeigler-Hill and T. K. Shackelford (Cham: Springer), 5011–5015.

Hall, J. A., Andrzejewski, S. A., and Yopchick, J. E. (2009). Psychosocial correlates of interpersonal sensitivity: a meta-analysis. *J. Nonverbal Behav.* 33, 149–180. doi: 10.1007/s10919-009-0070-5

Hancock, J.T., Liu, S.X., French, M., and Mieczkowski, H., (2019). Social media use and psychological well-being: a meta-analysis. In Proceeding of the 69th annual international communication association conference (ICA), Washington, DC, 24–28.

Hänggi, Y. (2004). Stress and emotion recognition: an internet experiment using stress induction. *Swiss J. Psychol.* 63, 113–125. doi: 10.1024/1421-0185.63.2.113

Hayes, A. F., and Rockwood, N. J. (2017). Regression-based statistical mediation and moderation analysis in clinical research: observations, recommendations, and implementation. *Behav. Res. Ther.* 98, 39–57. doi: 10.1016/j.brat.2016.11.001

Ivie, P. A., Moses, L. J., and Allen, N. B. (2020). A meta-analysis of the association between adolescent social media use and depressive symptoms. *J. Affect. Disord.* 275, 165–174. doi: 10.1016/j.jad.2020.06.014

Jones, S. M., and Heerey, E. A. (2022). Co-rumination in social networks. *Emerg. Adulthood* 10, 1345–1360. doi: 10.1177/21676968221111316

Keiser, B. E. (1987). Practical competitor intelligence. *Plan. Rev.* 15, 14–45. doi: 10.1108/eb054200

Keles, B., McCrae, N., and Grealish, A. (2020). A systematic review: the influence of social media on depression, anxiety and psychological distress in adolescents. *Int. J. Adolesc. Youth* 25, 79–93. doi: 10.1080/02673843.2019.1590851

Kiesler, S., Siegel, J., and McGuire, T. W. (1984). Social psychological aspects of computer-mediated communication. *Am. Psychol.* 39, 1123–1134. doi: 10.1037/0003-066x.39.10.1123

Krasnova, H., Wenninger, H., Widjaja, T., and Buxmann, P. (2013). *Envy on Facebook: A hidden threat to users' life satisfaction? 11th international conference on Wirtschaftsinformatik, 27th February – 01st march 2013*, Leipzig, Germany

Kraut, R., Patterson, M., Lundmark, V., Kiesler, S., Mukophadhyay, T., and Scherlis, W. (1998). Internet paradox: a social technology that reduces social involvement and psychological well-being? *Am. Psychol.* 53, 1017–1031. doi: 10.1037/0003-066x.53.9.1017

Kross, V. P., Sheppes, G., Costello, C. K., Jonides, J., and Ybarra, O. (2021). Social media and well-being: pitfalls, Progress, and next steps. *Trends Cogn. Sci.* 25, 55–66. doi: 10.1016/j.tics.2020.10.005

Kross, E., Verduyn, P., Demiralp, E., Park, J., Lee, D. S., Lin, N., et al. (2013). Facebook use predicts declines in subjective well-being in young adults. *PLoS One* 8:e69841. doi: 10.1371/journal.pone.0069841

Lawson, K. M., and Robins, R. W. (2021). Sibling constructs: what are they, why do they matter, and how should you handle them? *Personal. Soc. Psychol. Rev.* 25, 344–366. doi: 10.1177/10888683211047101

Lin, L., Yi, L., Sidani, J. E., Shensa, A., Radovic, A., Miller, E., et al. (2016). Association between social media use and depression among U.S. young adults. *Depress. Anxiety* 33, 323–331. doi: 10.1002/da.22466

Liu, D., Baumeister, R. F., Yang, C. C., and Hu, B. (2019). Digital communication media use and psychological well-being: a meta-analysis. *J. Comput.-Mediat. Commun.* 24, 259–273. doi: 10.1093/jcmc/zmz013

Lovibond, P. F., and Lovibond, S. H. (1995). The structure of negative emotional states: comparison of the depression anxiety stress scales (DASS) with the Beck depression and anxiety inventories. *Behav. Res. Ther.* 33, 335–343. doi: 10.1016/0005-7967(94)00075-u

McCredie, M. N., and Morey, L. C. (2019). Who are the Turkers? A characterization of MTurk workers using the personality assessment inventory. *Assessment* 26, 759–766. doi: 10.1177/1073191118760709

Meier, A., Gilbert, A., Börner, S., and Possler, D. (2020). Instagram inspiration: how upward comparison on social network sites can contribute to well-being. *J. Commun.* 70, 721–743. doi: 10.1093/joc/jqaa025

Meier, A., and Krause, H. V. (2022). Does passive social media use harm well-being? *J. Media Psychol.* doi: 10.1027/1864-1105/a000358

Meier, A., and Schäfer, S. (2018). The positive side of social comparison on social network sites: how envy can drive inspiration on Instagram. *Cyberpsychol. Behav. Soc. Netw.* 21, 411–417. doi: 10.1089/cyber.2017.0708

Nie, N. (2001). Sociability, interpersonal relations, and the internet. *Am. Behav. Sci.* 45, 420–435. doi: 10.1177/00027640121957277

Noon, E. J., and Meier, A. (2019). Inspired by friends: Adolescents' network homophily moderates the relationship between social comparison, envy, and inspiration on Instagram. *Cyberpsychol. Behav. Soc. Netw.* 22, 787–793. doi: 10.1089/cyber.2019.0412

Nowicki, S., and Duke, M. P. (1994). Individual differences in the nonverbal communication of affect: the diagnostic analysis of nonverbal accuracy scale. *J. Nonverbal Behav.* 18, 9–35. doi: 10.1007/bf02169077

O'Day, E. B., and Heimberg, R. G. (2021). Social media use, social anxiety, and loneliness: a systematic review. *Comput. Hum. Behav. Rep.* 3:100070. doi: 10.1016/j.chbr.2021.100070

Pereira-Lima, K., and Loureiro, S. R. (2015). Burnout, anxiety, depression, and social skills in medical residents. *Psychol. Health Med.* 20, 353–362. doi: 10.1080/13548506.2014.936889

Reer, F., Tang, W. Y., and Quandt, T. (2019). Psychosocial well-being and social media engagement: the mediating roles of social comparison orientation and fear of missing out. *New Media Soc.* 21, 1486–1505. doi: 10.1177/1461444818823719

Ruben, M. A., Hall, J. A., Curtin, E. M., Blanch-Hartigan, D., and Ship, A. N. (2015). Discussion increases efficacy when training accurate perception of patients' affect. *J. Appl. Soc. Psychol.* 45, 355–362. doi: 10.1111/jasp.12301

Ruben, M. A., Stosic, M. D., Correale, J., and Blanch-Hartigan, D. (2021). Is technology enhancing or hindering interpersonal communication? A framework and preliminary results to examine the relationship between technology use and nonverbal decoding skill. *Front. Psychol.* 11. doi: 10.3389/fpsyg.2020.611670

Santarossa, S., and Woodruff, S. J. (2017). Exploring the relationship of social networking sites on body image, self-esteem, and eating disorders. *Social Media Soc* 3:205630511770440. doi: 10.1177/2056305117704407

Schlegel, K., Boone, R. T., and Hall, J. A. (2017). Individual differences in interpersonal accuracy: a multi-level meta-analysis to assess whether judging other people is one skill or many. *J. Nonverbal Behav.* 41, 103–137. doi: 10.1007/s10919-017-0249-0

Schlegel, K., Grandjean, D., and Scherer, K. R. (2014). Introducing the Geneva emotion recognition test: an example of Rasch-based test development. *Psychol. Assess.* 26, 666–672. doi: 10.1037/a0035246

Schlegel, K., and Scherer, K. R. (2016). Introducing a short version of the Geneva emotion recognition test (Gert-S): psychometric properties and construct validation. *Behav. Res. Methods* 48, 1383–1392. doi: 10.3758/s13428-015-0646-4

Shapiro, D. N., Chandler, J., and Mueller, P. A. (2013). Using mechanical Turk to study clinical populations. *Clin. Psychol. Sci.* 1, 213–220. doi: 10.1177/2167702612469015

Siegel, J., Dubrovsky, V., Kiesler, S., and McGuire, T. W. (1986). Group processes in computer-mediated communication. *Organ. Behav. Hum. Decis. Process.* 37, 157–187. doi: 10.1016/0749-5978(86)90050-6

Swar, B., and Hameed, T. (2017). "Fear of missing out, social media engagement, smartphone addiction and distraction: moderating role of self-help mobile apps-based interventions in the youth" in *In Proceedings of the 10th international joint conference on biomedical engineering systems and technologies*

Trifiro, , and Gerson, J. (2019). Social media usage patterns: research note regarding the lack of universal validated measures for active and passive use. *Social Media Society* 5:205630511984874. doi: 10.1177/2056305119848743

Vadukul, A. (2022). 'Luddite' teens don't want your likes when the only thing better than a flip phone is no phone at all. New York Times. Available at: https://www.nytimes.com/2022/12/15/style/teens-social-media.html

Vahedi, Z., and Zannella, L. (2021). The association between self-reported depressive symptoms and the use of social networking sites (SNS): a meta-analysis. *Curr Psychol* 40, 2174–2189. doi: 10.1007/s12144-019-0150-6

Valkenburg, P. M., Beyens, I., Pouwels, J. L., van Driel, I. I., and Keijsers, L. (2022a). Social media browsing and adolescent well-being: challenging the "passive social media use hypothesis". *J. Comput.-Mediat. Commun.* 27:zmab015. doi: 10.1093/jcmc/zmab015

Valkenburg, P. M., van Driel, I. I., and Beyens, I. (2022b). The associations of active and passive social media use with well-being: a critical scoping review. *New Media Soc.* 24, 530–549. doi: 10.1177/14614448211065425

Verduyn, P., Gugushvili, N., and Kross, E. (2022). Do social networking sites influence well-being? The extended active-passive model. *Curr. Dir. Psychol. Sci.* 31, 62–68. doi: 10.1177/09637214211053637

Verduyn, G. N., Massar, K., Täht, K., and Kross, E. (2020). Social comparison on social networking sites. *Curr. Opin. Psychol.* 36, 32–37. doi: 10.1016/j.copsyc.2020.04.002

Verduyn, L. D. S., Park, J., Shablack, H., Orvell, A., Bayer, J., Ybarra, O., et al. (2015). Passive Facebook usage undermines affective well-being: experimental and longitudinal evidence. *J. Exp. Psychol. Gen.* 144, 480–488. doi: 10.1037/xge0000057

Verduyn, P., Ybarra, O., Résibois, M., Jonides, J., and Kross, E. (2017). Do social network sites enhance or undermine subjective well-being? A critical review. *Soc. Issues Policy Rev.* 11, 274–302. doi: 10.1111/sipr.12033

Walther, J. B., and Burgoon, J. K. (1992). Relational communication in computer-mediated interaction. *Hum. Commun. Res.* 19, 50–88. doi: 10.1111/j.1468-2958.1992.tb00295.x

Williams, B. T., and Gray, K. M. (2013). The relationship between emotion recognition ability and social skills in young children with autism. *Autism* 17, 762–768. doi: 10.1177/1362361312465355

WorkHuman (2019), "The future of work is human: findings from the work", Human Analytics & Research Institute Survey. Available at: www.workhuman.com/press-releases/White_Paper_The_Future_of_Work_is_ Human.pdf (Accessed November 12, 2019).

Wright, K. B., Rosenberg, J., Egbert, N., Ploeger, N. A., Bernard, D. R., and King, S. (2013). Communication competence, social support, and depression among college students: a model of Facebook and face-to-face support network influence. *J. Health Commun.* 18, 41–57. doi: 10.1080/10810730.2012.688250

Yao, M. Z., and Zhong, Z.-J. (2014). Loneliness, social contacts and internet addiction: a cross-lagged panel study. *Comput. Hum. Behav.* 30, 164–170. doi: 10.1016/j.chb.2013.08.007

Yin, X. Q., de Vries, D. A., Gentile, D. A., and Wang, J. L. (2019). Cultural background and measurement of usage moderate the association between social networking sites (SNSs) usage and mental health: a meta-analysis. *Soc. Sci. Comput. Rev.* 37, 631–648. doi: 10.1177/0894439318784908

Yoon, S., Kleinman, M., Mertz, J., and Brannick, M. (2019). Is social network site usage related to depression? A meta-analysis of Facebook–depression relations. *J. Affect. Disord.* 248, 65–72. doi: 10.1016/j.jad.2019.01.026

frontiers | Frontiers in Computer Science

TYPE Original Research
PUBLISHED 13 July 2023
DOI 10.3389/fcomp.2023.972186

OPEN ACCESS

EDITED BY
Kostas Karpouzis,
Panteion University, Greece

REVIEWED BY
Ahmad Rabiah,
Technical University of Malaysia Malacca, Malaysia
Olga Angelopoulou,
University of Warwick, United Kingdom

*CORRESPONDENCE
Julian Romeo Hildebrandt
✉ hildebrandt@comm.rwth-aachen.de

RECEIVED 17 June 2022
ACCEPTED 22 June 2023
PUBLISHED 13 July 2023

CITATION
Hildebrandt JR, Schomakers E-M, Ziefle M and Calero Valdez A (2023) Understanding indirect users' privacy concerns in mobile forensics — A mixed method conjoint approach.
Front. Comput. Sci. 5:972186.
doi: 10.3389/fcomp.2023.972186

Understanding indirect users' privacy concerns in mobile forensics — A mixed method conjoint approach

Julian Romeo Hildebrandt[1]*, Eva-Maria Schomakers[1], Martina Ziefle[1] and André Calero Valdez[2]

[1]Chair of Communication Science, Human-Computer Interaction Center, RWTH Aachen University, Aachen, Germany, [2]Human-Computer Interaction and Usable Safety Engineering, Institute for Multimedia and Interactive Systems, Universität zu Lübeck, Lübeck, Germany

Introduction: The advent of mobile forensics toolkits introduces a technological leverage that allows legal authorities to access information stored on mobile devices, thus shining a new light on law enforcement procedures. These pocket-sized devices, mobile phones, accumulate a plethora of user data, effectively becoming a beacon for individual identification. However, the prospect of exploring this data within a criminal inquiry raises palpable concerns about potential privacy encroachments. Consequently, there exists an urgent need to balance the instrumental value of these technologies with their potential to intrude upon privacy, ensuring a framework that remains legally and ethically sound.

Methods: In our study, we offer a contemplative view on the public reception of such measures, informed by interviews and a conjoint study conducted across two representative cohorts from Germany and Austria (n = 2040).

Results: Our analysis indicates a marked preference for the release of geo-spatial data over more personal content, such as photographs. Additionally, respondents showed a higher acceptance for automated analysis in comparison to human manual evaluation. The divergence between the two countries was negligible.

Discussion: In summary, despite the inherent concerns, the use of these mobile forensics tools demonstrated a high degree of public acceptance. The results highlight the significance of aligning legislation based on data types rather than analysis purposes, which can enhance the general public's comprehension of laws and potentially contribute to societal advancement. Furthermore, the research emphasizes the importance of ethical evaluations and transparent communication in the implementation of automated mobile forensics systems for civil security purposes, effectively addressing concerns regarding privacy infringement and data analysis.

KEYWORDS

privacy, mobile forensics, conjoint, mixed method, acceptance

1. Introduction

There are enormous societal benefits associated with the increasing global collection and use of digital cloud data (Bryant et al., 2008; Gudivada et al., 2015). Amidst the demands for an effective and prosperous digital transformation, societies find themselves confronted with a crucial trade-off between comprehensive data collection and utilization for the greater good, on one hand, and the potential risks of data misuse on multiple levels, on the other (Dritsas et al., 2006; Anandaraj and Kemal, 2017; Jindal et al., 2023). Among the benefits, social and societal progress in all areas, e.g., health, commerce, mobility, production, security, and safety, has been reached and will still further develop through the availability of Big Data. However, the use of ubiquitous data comes with serious legal, technical, and social

disadvantages, mainly in terms of privacy loss, intrusion into intimacy, and personal rights (Nissenbaum, 2010; Hayes et al., 2020; Huang et al., 2023). Hence, as a side effect of the digital transformation in general, digital information and communication technology devices have become not only carriers of sensitive data but also more and more unique to their user. Based on these characteristics, a smartphone could be used as a personal identifier and/or a carrier of evidence in a law enforcement scenario. Equipped with mobile forensic systems, law authorities can collect, examine, analyze, and report digital evidence without damaging the device (Dogan and Akbal, 2017; Losavio et al., 2018).

While this area of digital mobile forensics is mostly covered by the study of technical perspectives and possibilities, public acceptance and the informed decision for data sharing and its consequences are equally important. Hence, questions of social, legal, and ethical implications have to be discussed to ensure adequate democratic coverage (Gantz and Reinsel, 2012; Dhirani et al., 2023). Digital mobile forensics is an emerging field that specifically addresses processes, methods, and analytics of any material that can be collected in digital devices (Dritsas et al., 2006; Cruz-Cunha and Mateus-Coelho, 2020; Alenezi, 2023). Digital forensics is mostly related to the use of mobile device data in the context of crime prosecution (Carrier, 2003; Du et al., 2020; Saranya and Usha, 2023). Criminal application fields cover breaking the law in the context of data collection and analysis and use and involve legislative and executive entities but also civil cases in the context of privacy protection and personal rights.

The discourse around digital mobile forensics should not be confined to technical possibilities and capacities. The issue of public acceptance plays a substantial role and deserves equal consideration. Thus, it is crucial that we engage in a discourse around the social, legal, and ethical implications to ensure that democratic principles are upheld in the deployment of these technologies. As privacy is a fundamental right (Pfisterer, 2019) and information privacy comprises the control over (the access and use of) personal information (Ermakova et al., 2014), dilemmas may arise between the individual right to privacy and various benefits of mobile forensics for law enforcement and public security, such as improved evidence provision (Al-Dhaqm et al., 2020) or detection of suspicious patterns (Barmpatsalou et al., 2018).

Mobile phones inadvertently collect an extensive array of data without the user's awareness or control over the data collection process, such as dating apps analyzing music preferences (Hayes et al., 2020). Ordinarily, these data are not shared with third parties. However, in law enforcement scenarios, such data may be employed. This distinguishes mobile forensics from other privacy-sensitive domains such as social networks or smart homes, where users intentionally disclose personal information to data providers. Furthermore, there is insufficient research examining the evaluation of these data by indirect users who do not directly interact with the system. It is important to differentiate between direct and indirect users, with direct users engaging in direct interactions with the system while indirect users are affected by its use, even without direct interaction (ISO, 2011). Indirect users, in this context, are the owners of phones that are being analyzed by criminal investigators, as owners of smartphones never interact with the mobile forensic software directly. Another example would be uninvolved peers of suspects—whose mobile phones were confiscated—as data of the peers might be on that phone.

The preservation of privacy is a fundamental right (Pfisterer, 2019). Control over personal information, which forms the cornerstone of information privacy according to the study by Ermakova et al. (2014), presents an intriguing quandary when juxtaposed with the advantages of mobile forensics. This technology equips law enforcement and public security agencies with enhanced capabilities, including effective evidence collection (Al-Dhaqm et al., 2020) and the identification of suspicious patterns (Barmpatsalou et al., 2018). These benefits are made possible by the copious amount of data that mobile phones amass as part of their regular operation, often without the user's knowledge or consent. Consequently, there is an urgent need for improved policymaking (Hayes et al., 2020). It is important to note that, in normal circumstances, this information remains isolated and is not shared with third parties.

However, the narrative changes dramatically in a law enforcement scenario where this data may be actively exploited. This situates mobile forensics distinctly apart from other privacy-sensitive spheres such as social networks or smart homes, where personal data are knowingly and willingly shared with service providers. What remains inadequately explored is the evaluation of these systems, particularly from the perspective of "indirect users," who may never interact with the system directly, yet find themselves affected by its use (ISO, 2011). This distinction between direct and indirect users unveils an important layer of user interaction and its implications for privacy and security demand further investigation.

2. Questions addressed and empirical procedure

We aim to provide first insights into the public's acceptance of and preferences for mobile forensics. To do so, we use a two-step empirical research approach. In the first exploratory step, we explore the public's opinion on mobile forensics and important influencing factors using guided interviews. In the second step, we apply a choice-based conjoint approach to experimentally assess the importance of relevant influencing factors based on large, census-representative samples from Germany (n = 1,039) and Austria (n = 1,001). The selection of countries refers to a joint project in which both countries were involved.

The joint project SmartIdentifikation aimed to explore the feasibility of utilizing smartphone-based or border-collected data in a socially accepted, ethically sound, and legally secure manner for analyzing refugee flows, detecting smuggling routes, and identifying involved individuals [Federal Ministry of Education and Research , BMBF]. By employing a two-country setting, we were able to shed light on potential international police cooperation, a vital element in investigating crimes such as human trafficking (Marturana et al., 2011). Additionally, this approach allowed us to address the research gap concerning comparative analysis (Kimmelmann et al., 2022) among states already engaged in border and migration control cooperation (Karamanidou et al., 2020; Comte and Lavenex, 2022), thus enhancing our understanding of factors that influence civil society participation in the context of refugee migration and integration (Simsa, 2017).

Our study primarily focused on exploring the public perception and acceptance of mobile forensics. We aimed to identify the key factors driving public acceptance of such technologies and determine which types of data, levels of automation, and analysis purposes were most widely accepted.

We thereby contribute to an understanding of the public's acceptance of mobile forensics as well as privacy preferences of people, whose data may be used but who are not in control over its usage, e.g., contact persons of crime suspects. The results contribute to a better understanding of public privacy perception and provide implications for developers and investigators.

3. Materials and methods

In the following, we describe our two-step consecutive empirical research approach: an exploratory qualitative interview study and a choice-based conjoint (CBC) online survey.

3.1. Interview study

To gain a first understanding of the acceptance of mobile forensics use cases and identify factors to be included in the CBC study, interviews with German citizens of different age groups, gender, and (professional) background were conducted. The interview study was guided by the following research question:

- RQ1: What are the relevant factors driving the public acceptance of mobile forensics?

To answer this research question, we conducted semi-structured guideline-based interviews with a professional moderator. The average interview length was approximately 20 min and participants were allowed to abort the interview at any time. The topic of mobile forensics was introduced by a narrative scenario of two investigators talking about mobile forensics. Participants were asked about their previous knowledge and feelings about this technology, their conditionals and concerns for an accepted usage, and, in particular, about a sub-scenario where not only the smartphone of the subject is used for investigation but also the device of a close relative. A final question was asked to the interview participants about their general attitude and feelings again.

All interviews were anonymized and converted to text according to GAT2 basic transcription convention (Selting and Auer, 2011). Every transcript was chosen as a sampling unit and fully considered as a recording unit; short stand-alone responses as the content unit and detailed answers as the context unit (Krippendorff, 2018). Qualitative analysis was conducted as thematic qualitative text analysis by Kuckartz (2014). After the first step of initial text work, four deductive main categories were identified: *General attitude*, *Type of data*, *Conditionals*, and *Role of other people's data*.

Participants were recruited to cover different age groups, genders, and education levels. We aimed for a random sample of the general population in order to capture typical average behaviors of mobile phone users. Overall, $n = 9$ interviews were conducted with 4 female and 5 male participants of varying professions. Age ranged from $min = 24$ years to $max = 72$ years with an average of $M = 44.66 (SD = 16.40)$, while 3 participants were of higher, 4 of intermediate, and 2 of lower scholarly education levels. All participants were frequent smartphone users with no prior knowledge of mobile forensics. Participation was not gratified, and all participants volunteered to take part in the interview study.

TABLE 1 Attributes and levels included in the CBC.

Attributes	Levels
Type of data	Location data, sms/mms, messenger data (e.g. Whatsapp), device data, user account data, image data
Analysis automation level	Automated, manually
Analysis purpose	Prosecution, prevention

3.2. Choice-based conjoint (CBC) study

Building on the findings from the interview study, a CBC questionnaire was designed. The leading research questions for the conjoint study were as follows:

- RQ2: How important are the three factors such as type of data, automation level, and analysis purpose for public acceptance?
- RQ3: Which types of data, automation levels, and analysis purposes are most accepted?
- RQ4: Are there differences in these preferences between the German and Austrian public?

In CBC, complex decision processes can be mimicked as several factors are evaluated in conjoint and need to be weighed against each other, which provides a more realistic evaluation situation than typical rating scales (Sawtooth Software, Inc.). CBC is a decompositional procedure, meaning that participants choose their favorite from several presented options. From this, the relevance of individual *attributes* is derived based on the assumption that the preference for the overall product/option is a function of a set of explanatory variables, named attributes (Baier et al., 2009). In CBC, these attributes are represented by their associated *levels* which are experimentally varied across the presented options. For example, the attribute "color" would be represented by varying levels such as "green," "red," and "blue".

3.2.1. Selection of attributes

The selection of attributes and their associated levels, i.e., the operationalization of the variables, is critical as the results are relative to one another and are valid only for the used combination of attributes and levels. Therefore, the most relevant attributes of a decision-making process must be covered (Johnson and Orme, 2003). Because of the exploratory nature of our study, the attributes are selected based on the focus groups' findings, and the corresponding levels are developed to reflect mobile forensics realistically and be comprehensive for the participants. Table 1 displays the attributes and their associated levels.

As one essential factor for the acceptance of mobile forensics, the type of data was identified in the focus groups. The type of data—and its individually perceived sensitivity—is one of the main influences on the acceptance of data collection and analysis, which is also confirmed by empirical research in other contexts, in which the data provider intentionally discloses the data (Li, 2011; Mothersbaugh et al., 2012; Schomakers et al., 2020). As the perceived sensitivity of data increases, individuals tend to have higher privacy concerns regarding its usage, which consequently leads to a lower willingness to disclose that data (Bansal et al., 2010; Mothersbaugh et al., 2012). Previous research has indicated that the perception of data sensitivity is influenced by various factors. For instance, the sensitivity of data can be influenced by whether it contains personally identifying information (Malheiros et al., 2013), is associated with physical, monetary, social, or psychological risks (Milne et al., 2017), and originates from a particular source (Rohm and Milne, 2004). Additionally, the perception of information sensitivity is highly individual and dependent on culture (Markos et al., 2017; Schomakers et al., 2019). Location data have been highlighted as very sensitive data (besides medical and financial data which are not included here) (Staiano et al., 2014; Milne et al., 2017; Schomakers et al., 2019, 2020). As privacy and privacy-related preferences are also strongly dependent on the context (Nissenbaum, 2010; Acquisti et al., 2015; Schomakers et al., 2021b), an assessment of the acceptance of data types for mobile forensics is still needed.

Based on previous research, we hypothesize that:

- H1: The analysis of location data is less accepted than the analysis of the other types of data. The hypothesis is based on previous findings from the literature which suggest that location data are perceived as highly sensible and intrusive from the users' point of view (Staiano et al., 2014; Milne et al., 2017; Schomakers and Ziefle, 2019; Schomakers et al., 2020).

Furthermore, the interview results shed light on the privacy invasion by the person who accesses and investigates the data. However, mobile forensics does not necessarily require a human investigator manually looking at the data; also automated data analysis using artificial intelligence is possible (Sikos, 2021). Therefore, it is highly interesting to gain insights into the public's preferences for either manual or automated data analysis. As the interview findings suggest that one of the key privacy concerns for mobile forensics is the *human* investigator looking at personal data, automated data analysis may reduce the perceived privacy intrusion. Hypothesis 2 is, therefore, related to the qualitative study which has been carried out prior to the Choice-Based Conjoint study. Hypothesis H2 reads:

- H2: Automated data analysis is preferred to manual data analysis. This hypothesis is based on our qualitative findings (see Section 4.1).

We identified the increased civil security as well as the severity of the crime to be additional conditionals for the acceptance of mobile forensics. These factors describe the purposes of the analysis and refer, on the one hand, to the societal benefit of the data analysis, and, on the other hand, to the legitimacy of the data analysis. The importance of the gained (individual and/or societal) benefit for privacy decisions has been emphasized in various studies (Dinev et al., 2006; Calero Valdez and Ziefle, 2019; Schomakers et al., 2021b), and the purpose of data collection and analysis is of high relevance for privacy rights (Voigt and Von dem Bussche, 2017; Jasserand, 2018).

A relevant research duty for democratic coverage is to gain insights into the public's acceptance of mobile forensics for different analysis purposes. Particularly, the differentiation of crime prevention vs. crime prosecution is highly relevant as it is an important differentiation for law enforcement bodies and their legal rights. However, empirical data about the public's preferences regarding the use of mobile forensics in these cases are lacking.

From a benefit perspective, an even higher benefit may arise from the prevention of planned crimes in contrast to the prosecution of past crimes. However, the legitimacy of investigating the data of persons who may plan to commit a crime may be controversial—as elegantly discussed in the Hollywood movie Minority Report. Therefore, we study the public preferences regarding the analysis purpose, distinguishing crime prosecution and crime prevention. Considering our interview results, the barriers of crime prevention are hypothetically higher because of negative narratives (e.g., "surveillance state").

- H3: Mobile forensics for crime prosecution is more accepted than crime prevention. This hypothesis is based on our qualitative findings (see Section 4.1).

3.2.2. The study design

The choice tasks were embedded into a questionnaire, which introduced the topic of mobile forensics and assessed sociodemographic characteristics as well as key attitudes of the respondents. The questionnaire started with questioning age, gender, education, and profession. In the second part, technology readiness (Neyer et al., 2012) and disposition to value privacy (Xu et al., 2008) were addressed using six-point Likert scale. The final question of the survey was a single net promoter score item (1–10) whether participants felt positive or negative about the use of mobile forensics.

For the CBC part, the participants were asked to put themselves in the following scenario: *Imagine that a person suspected of being connected to a crime has been picked up and is carrying a smartphone. This smartphone has been confiscated as evidence and is now to be evaluated to preserve evidence.* Then, 10 choice tasks were carried out. In each choice task, the participants were presented with two options comprised of the three attributes, such as type of data (what), analysis automation level (how), and analysis purpose (why) with varying levels (cf. Figure 1). The task was to choose the most acceptable option.

3.2.3. Recruiting and data analysis

The survey data were gathered from panel members of an independent research company. Quotas were set to acquire a

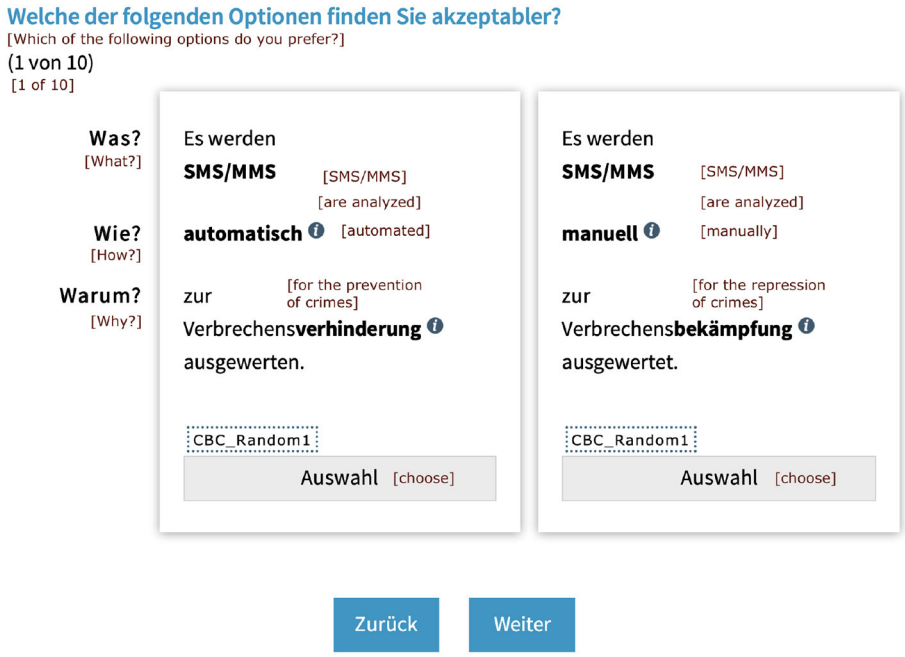


FIGURE 1
Example of a choice task (translations are added in red color for the publication).

census-representative sample regarding age, gender, federal state, and education level of adults from Austria and Germany. The quality of the data was ensured using several attention questions and excluding speeders as well as incomplete data sets.

The conjoint data were analyzed using Hierarchical Bayes (HB) estimation to assess individual-level values for the *relative importance* of the attributes and the *part-worth utilities* of the levels. The relative importance shows how important each attribute is for the acceptance of mobile forensics. The part-worth utilities describe the preferences for the levels. These are reported as zero-centered differences to allow comparisons between the levels of the different attributes. The mean values for the German and Austrian subsamples are calculated and reported separately. The mean root likelihood of the model is 0.65. As two concepts are shown per choice task, the worst would be 0.5 so that the estimates fit the model moderately.

To test differences between the Austrian and German samples, *t*-tests were used in the sample description. For the conjoint data, Bayesian *t*-tests were used on the individual zero-centered utility scores and the individual relative importance. We report the *t*-test results alongside the Bayes factor as these are more commonly known, but we interpret mainly the Bayes factor. The level of significance was set at 5%. To avoid alpha error inflation, Bonferroni–Holm correction is used on the significance values when several tests were calculated.

3.2.4. Sample description

The sample comprises 1,001 Austrian and 1,038 German participants. The demographic characteristics of the sample are displayed in Table 2 for the overall sample and the national subsamples.

The technology readiness of the sample is rather high ($M = 4.40$, $SD = 0.87$) and shows no significant differences between the German and Austrian samples [$t_{(2,037)} = 1.30$, $p = 0.193$]. The privacy disposition is rather high as well with $M = 4.01$ ($SD = 1.11$) and shows no significant differences between the nations [$t_{(2,037)} = 1.91$, $p = 0.056$].

4. Results

4.1. Qualitative interview results

Figure 2 shows the final result of the thematic content analysis, with *Challenges* being an inductive main category. Overall, we could identify 156 content units and 16 inductive sub-categories.

While asking participants about their *general attitude* toward the provided mobile forensics scenario, most participants expressed a positive or even enthusiastic position, as far as some conditions were met. Only a few participants were rather skeptical, and none was fundamentally against mobile forensics.

As already mentioned, participants were not explicitly asked about the types of data that could be used for investigation. They mentioned *local data*, such as GPS or mobile network, text message data, that are stored directly on the device (*SMS / MMS*), and *messenger data*, that are stored in apps or servers, *device data* such as contacts or phone identifiers, as well as *user account data* for social networks or online forums. All of these types of data were mentioned without a specific connotation in terms of privacy but more under the perspective of data variety. *Image data*, such as

pictures or camera images, were described as quite privacy intrusive or even intimate.

When talking about conditionals for mobile forensic systems, participants stated that the system should lead to *increased civil security* by providing clear evidence and faster investigation. In addition, mobile forensics might also provide a feeling of security through crime prevention, either by bringing criminals to justice and therefore preventing further crimes, but also by preventing terrorist attacks.

While discussing the *privacy invasion* of the system, participants implicitly tied the invasiveness to human primary users. They argued that they might feel uncomfortable thinking about other people investigating their smartphones and noticed that policemen (not police as an institution) could see everything they want. One participant even argued that policemen should put under surveillance while using the system to prevent privacy intrusion.

Data security was found to be a major conditional in a way that only the executing police should be able to use the data for an explicit investigation. Finally, subjects disagreed on whether the system should be used for particularly serious crimes, such as murder, or also for lighter ones, such as vandalism.

When being asked about the *role of other people's data*, some participants noticed during the interview that most types of data carry information about non-suspected individuals, i.e., message data always contains information about at least one other individual. In addition to that, participants mostly argued that the analysis of suspects' smartphones and data—with the already mentioned characteristics—should be the standard case of mobile forensics. On the other hand, the analysis of the *data of suspects' contact persons* was considered potentially useful or even generally accepted. One participant raised the condition that the smartphone of contact persons should be provided voluntarily, and another participant mentioned that close relatives of subjects might have the right to refuse to testify, which should be adapted to this digital context. Interestingly, when discussing the data of completely uninvolved people, some participants switched their perspectives and ensured that they would provide their own smartphone data to help the police. This was explained by the fact that this does not cause any harm to the participant and could furthermore contribute to the fight against crime.

The last main category contains upcoming *challenges*. Two participants mentioned that the benefit of investigation might arise only from the aggregation and correlation of data, such as people being in the same place on a regular basis. Furthermore, a mobile forensic system might even increase the effort for investigators by providing a huge amount of data or organizational effort, and finally, it was mentioned that *data manipulation* might provide an alibi to a criminal.

TABLE 2 Descriptive statistics for attribute importance scores.

	Total	Germany	Austria
Age	42.4 (SD = 12.8)	43.1 (SD = 13.0)	41.6 (SD = 12.6)
Gender			
Women	52.7%	52.0%	53.3%
Men	47.3%	48.0%	46.7%
Education			
No	0.7%	0.1%	1.3%
Basic secondary education	14.7%	9.4%	20.1%
Intermediate secondary education	15.1%	23.6%	6.3%
Apprenticeship	20.3%	20.9%	19.6%
High-school diploma	15.9%	13.5%	18.2%
University degree	30.1%	30.1%	30.2%
PhD or higher	3.3%	2.2%	4.4%
Employment			
Unemployed	13.5%	13.4%	13.7%
Employed	62.2%	63.3%	61.1%
Self-employed	8.6%	7.7%	9.6%
Public employment	3.8%	3.2%	4.5%
Retired	11.8%	12.4%	11.1%
Technology readiness	4.40 (SD = 0.87)	4.38 (SD = 0.84)	4.43 (SD = 0.91)
Privacy disposition	4.01 (SD = 1.11)	4.06 (SD = 1.02)	3.96 (SD = 1.19)

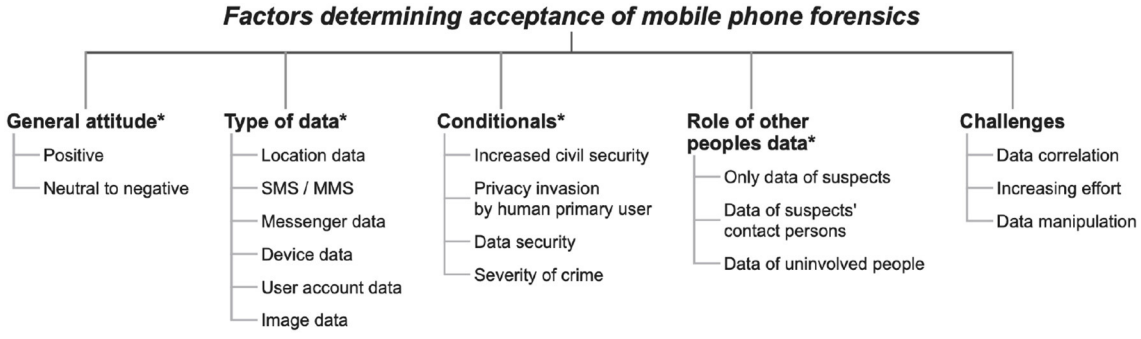


FIGURE 2
Final main category system. Asterisk denotes deductive categories.

4.2. Conjoint study results

Table 3 provides descriptive statistics for attribute importance scores for Germany, Austria, and both countries combined. Overall, the type of data was of the highest relative importance ($M = 48.54$, $SD = 17.86$), followed by the type of analysis ($M = 31.48$, $SD = 18.86$) and analysis purpose ($M = 19.97$, $SD = 15.34$).

The following bar chart (Figure 3) shows that the aforementioned order of average stays the same for German and Austrian subsamples. The inferential comparison yields a significant difference in the importance of the analysis purpose [$t_{(2,037)} = 2.83, adj.p = 0.014^{*}; BF_{10} = 2.65, err.\% < 0.001$], with analysis purpose being more important in Austria ($M = 20.96, SD = 16.03$) than in Germany ($M = 19.03, SD = 14.67$). However, this difference has a small effect size ($d = 0.13$).

Accordingly, Table 4 shows descriptive statistics and comparison for part-worth utilities. The most accepted type of data to be analyzed is location data ($M = 62.16, SD = 43.31$), followed by sms/mms data ($M = 4.03, SD = 38.00$). Data analysis, messenger data ($M = -3.06, SD = 38.72$), device data ($M = -6.88, SD = 56.90$), user account data ($M = -23.67, SD = 40.30$), and image data ($M = -32.59, SD = 43.07$) are less preferred. Regarding the other two attributes, automated analysis is favored over manually ($M = 18.04, SD = 52.03, resp., M = -18.04, SD = 52.03$), and analysis for prosecution is favored over prevention ($M = 15.85, SD = 34.33, resp., M = -15.85, SD = 34.33$).

Comparing Germany and Austria, there are four significant differences in the part-worth utilities of different data types. Location data [$t_{(2,037)} = 4.84, adj.p < 0.001^{***}; BF_{10} = 5141.46, err.\% < 0.001$] are slightly ($d = 0.21$) more favored in Austria ($M = 66.91, SD = 37.74$) than in Germany ($M = 57.69, SD = 47.56$). The significant difference regarding sms/mms [$t_{(2,037)} = 4.62, adj.p < 0.001^{***}; BF_{10} = 1894.87, err.\% < 0.001$] is of similar effect size ($d = 0.20$) and in the same direction: the analysis of sms/mms is slightly more accepted in Austria ($M = 7.93, SD = 37.00$) compared with Germany ($M = 0.20, SD = 38.51$). On the other hand, messenger data [$t_{(2,037)}$ =-3.69, $adj.p = 0.002^{**}; BF_{10} = 41.53, err.\% < 0.001$] and user account data [$t_{(2,037)}$ =-5.37, $adj.p < 0.001^{***}; BF_{10} = 75277.66, err.\% < 0.001$] are slightly more accepted in Germany (messenger data $M = -0.01, SD = 40.52$; user account data $M = -19.00, SD = 43.59$) than in Austria (messenger data $M = -6.30, SD = 36.43$; user account data $M = -28.53, SD = 35.97$). As with the previous ones, those effects are weak (messenger data $d = -0, 16$, user account data $d = -0.20$).

Finally, regarding the analysis purpose, there is another weak ($d = 0.12$ resp. $d = -0.12$) yet significant effect [$t_{(2,037)} = 2.75, adj.p = 0.030^{*}; BF_{10} = 2.12, err.\% < 0.001$, resp. $t_{(2,037)} = -2.75, adj.p = 0.036^{*}; BF_{10} = 2.12, err.\% < 0.001$]. Data analysis with the aim of prosecution is slightly more accepted in Austria ($M = 17.97, SD = 35.27$) than in Germany ($M = 13.79, SD = 33.31$), respective prevention of crime is slightly more accepted in Germany ($M = -13.79, SD = 33.31$) than in Austria ($M = -17.97, SD = 35.27$).

In summary, the type of data is more important than the automation analysis level and the analysis purpose, and the most accepted mobile forensic system uses location data and operates automated for the purpose of crime prosecution. Overall, participants were positive about the use of mobile forensics as described in the questionnaire and as stated in the net promoter score item ($M = 7.25, SD = 2.45, Md = 8.00$).

5. Discussion

In our study, we employed a mixed-method approach, combining qualitative interviews and a conjoint study, to compare the importance of various attributes in the context of mobile forensics. Our primary research goals were two-fold: First, to identify the relevant factors influencing public acceptance of mobile forensics; and second, to gain a deeper understanding of the disparities in mobile forensic system characteristics and the divergent perceptions between individuals in Germany and Austria.

5.1. Key findings

Our findings revealed substantial disparities in the average importance assigned to different attributes, namely, data particles, type of analysis, and analysis purpose. Notably, we unearthed a significant discrepancy that highlights the type of data as the most influential factor in determining the acceptance of mobile forensic approaches, followed by the type of analysis, be it automated or manual, which carries approximately half the weight in decision-making processes. Conversely, the purpose of analysis emerged as the least influential factor.

Regarding the average importances, we observed only marginal distinctions between Austria and Germany, with minimal variations detected across a few levels.

The survey yielded a key finding that underscores the varying degrees of sensitivity associated with different types of data. Specifically, when it comes to geospatial data, which can potentially lead to the easy and unique identification of users, the acceptance of its use was found to be the highest. On the other hand, image data exhibited the lowest acceptance of use. An intriguing observation emerged when considering the preferred mode of analysis: Participants surprisingly favored analysis conducted by automated systems over analysis performed by manually investigating officers. Evidently, the act of having our data viewed by a human was perceived as a more invasive breach of privacy compared with examination by an automated system, despite the latter's ability to scrutinize our data in a more comprehensive manner. Additionally, the analysis of qualitative data shed light on the overall positive perception of mobile forensic systems. This positive perception, however, hinges on the condition that the system design adequately addresses both primary and indirect users' requirements.

Contrary to our initial hypothesis (H1), the acceptance of location data was unexpectedly high among the subjects. It appears that participants may not be fully aware of the potential invasiveness associated with location data. Alternatively, it is plausible that location data are perceived as particularly useful or beneficial in some way, leading to its greater acceptance. Conversely, the analysis of image data was perceived as invasive and intimate. This suggests that participants recognize the personal

TABLE 3 Descriptive statistics and Bayesian country comparison for attribute importance scores.

	Germany $M(SD)$	Austria $M(SD)$	$t_{(2037)}$	Adj. p	d	BF_{10}	Err.%
Type of data	49.15(18.02)	47.90(17.72)	−1.58	0.228	−0.07	0.17	0.001
Automation analysis level	31.81(19.19)	31.14(18.54)	−0.80	0.421	−0.04	0.074	0.003
Analysis purpose	19.03(14.67)	20.96(16.03)	2.83	0.014*	0.13	2.65	<0.001

*$p < 0.05$.

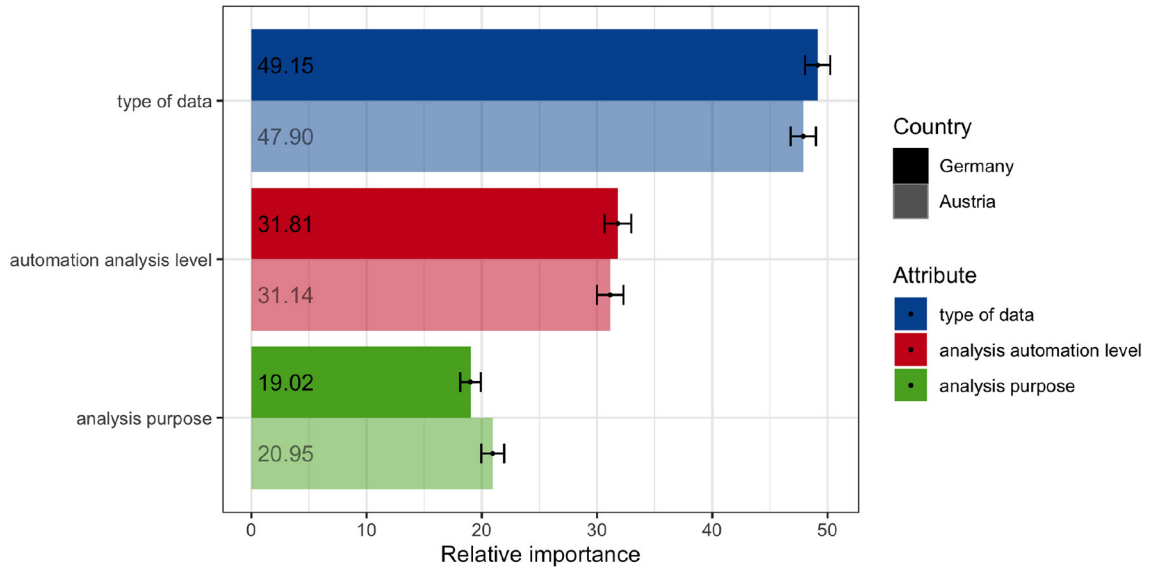


FIGURE 3
Average relative importance scores in bilateral comparison between Germany (n = 1,039) and Austria (n = 1,001). Error bars denote 95% CI, and labels denote the average score.

TABLE 4 Comparative descriptive statistics and Bayesian country comparison for part worth utility levels.

	Germany $M(SD)$	Austria $M(SD)$	$t_{(2037)}$	Adj. p	d	BF_{10}	Err.%
Location data	57.69 (47.56)	66.91 (37.74)	4.84	<0.001***	0.21	5141.46	<0.001
SMS/MMS	0.20 (38.51)	7.93 (37.00)	4.62	<0.001***	0.20	1894.87	<0.001
Messenger data	−0.01 (40.52)	−6.30 (36.43)	−3.69	0.002**	−0.16	41.53	<0.001
Device data	−5.81 (57.46)	−7.968 (56.34)	−0.85	0.789	−0.04	0.07	0.003
User account data	−19.00 (43.59)	−28.53 (35.97)	−5.37	<0.001***	−0.24	75277.66	<0.001
Image data	−33.07 (45.64)	−32.05 (40.24)	0.54	0.591	0.02	0.06	0.004
Automated	16.89 (53.13)	19.30 (50.84)	1.05	0.888	0.05	0.09	0.003
Manually	−16.89 (53.13)	−19.30 (50.84)	−1.05	1.184	−0.05	0.09	0.003
Prosecution	13.79 (33.31)	17.97 (35.27)	2.75	0.030*	0.12	2.12	<0.001
Prevention	−13.79 (33.31)	−17.97 (35.27)	−2.75	0.036*	−0.12	2.12	<0.001

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

and sensitive nature of image data, which could explain the lower acceptance of its use compared with other types of data.

Hypothesis H2 has been confirmed by the findings, as subjects exhibited a preference for automated analysis over manual analysis. This preference suggests that participants are more inclined toward the utilization of automated systems for evaluating data. Moreover, the perception that automated evaluation is less intrusive could be attributed to the absence of human involvement in directly inspecting, for example, images. Instead, the use of algorithms in the automated analysis might be perceived as less invasive. This observation supports the notion that users consider the level of effort exerted by investigators when evaluating the acceptability of

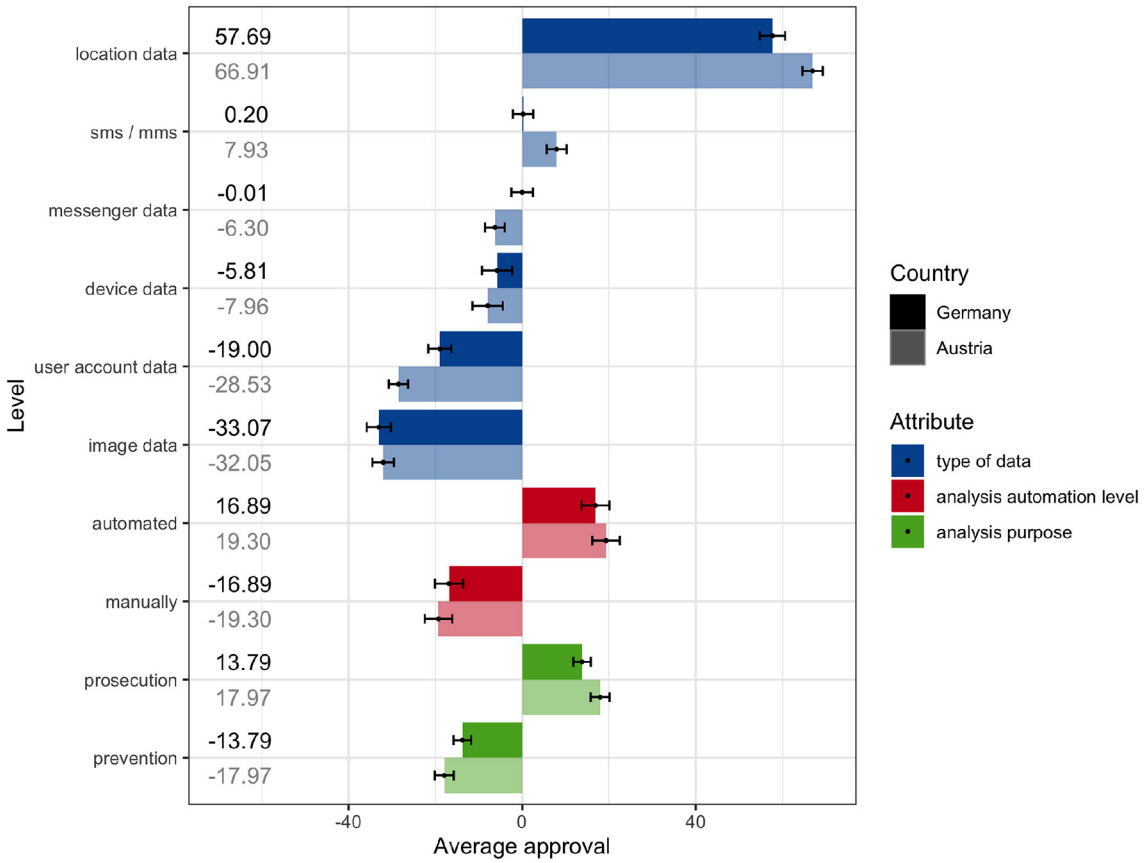


FIGURE 4
Average part-worth utilities in bilateral comparison between Germany (n = 1,039) and Austria (n = 1,001). Error bars denote 95% CI, and labels denote the average score.

data analysis methods. H3 could also be confirmed, as prosecution is more accepted than crime prevention. This might be due to dystopian connotation but also because of further and more in-depth conditionals.

Overall, we found only minor differences between Germany and Austria, in terms of privacy perceptions (Figure 4). The disparities in the sensitivity perception of different types of data between the two cultures also appear to be influenced by distinct privacy laws and practices, as highlighted by previous studies (Schomakers et al., 2019, 2021a). However, based on our acceptance study, there is little evidence to suggest that international cooperation on a bilateral level would be problematic, as the results demonstrate a high degree of similarity between the two countries. Comparatively, Germans tend to exhibit higher privacy concerns compared with many other countries, such as Turkey. Nevertheless, when comparing Germany to Austria, the differences in privacy concerns are relatively small (Krasnova and Veltri, 2010; Wilkowska et al., 2021). This similarity can potentially be attributed to cultural similarities, particularly the in-depth historical restrospection and reflection of World War II and subsequent political shifts. Furthermore, both Austria and Germany share a common ground in technology skepticism and media role (Metag and Marcinkowski, 2014).

5.2. Managerial recommendations and application

Several significant points emerge from the research that can provide valuable insights for various stakeholders and address societal issues.

5.2.1. Digital literacy and media education

First, there is a crucial need for enhanced digital literacy and media education. To make well-informed decisions concerning digital transformation, it is essential for both media and academic institutions to place greater emphasis on fostering digital literacy. A recent survey conducted in Europe (Eurostat, 2022) indicates varying levels of digital literacy across countries, highlighting the urgent necessity to educate citizens in all aspects of digital and mobile media usage (Ribble et al., 2004). Ribble et al. argue that digital education plays a pivotal role in promoting public awareness regarding the appropriate and intentional use of digital media, as well as a deep understanding of the consequences associated with digital behaviors, encompassing what is often referred to as Digital Citizenship.

"Digital citizenship can be defined as the norms of behavior with regard to technology use. As a way of understanding the

complexity of digital citizenship and the issues of technology use, abuse, and misuse, we have identified nine general areas of behavior that make up digital citizenship" (Ribble et al., 2004).

This claim includes a global call to action in media education and data literacy that includes knowledge about the potential dangers and benefits when using digital media (Rouvroy and Poullet, 2009; Ziefle et al., 2016). This claim is directed to public education and policymakers but is increasingly shifted to an individual responsibility of mobile device users (Rouvroy and Poullet, 2009; Tene and Polonetsky, 2012).

By doing so, the public can be better equipped to engage in discussions about digital transformation and make informed decisions based on a comprehensive understanding of the subject matter.

Based on the findings, we can offer the following recommendations to developers of mobile forensic systems: A system that gains acceptance from the general population should primarily operate in an automated manner, prioritizing the support of the primary user in their mission to contribute to civil security. Location data, while potentially beneficial in combating crime, present the least problematic aspect in this regard. There may even be potential for voluntary provision of such data, providing that robust measures are in place to mitigate any risks of data misuse.

In future, it will be crucial to determine the extent to which the analysis of image data continues to be regarded as intrusive, even in scenarios where human users do not have direct access to th image material.

5.2.2. Implications for legal authorities and policymakers

In light of the findings, it is advisable to recommend that legislative authorities align law formulations based on the types of data rather than the purpose of analysis. This approach would enhance the understandability and comprehensibility of laws for the general public. Such a shift could potentially mitigate the escalating tensions in the migration debate, which is often influenced by perceptions of the legality and enforcement of applicable laws. Moreover, this could indirectly contribute to the overall improvement of society by fostering civic engagement through voluntary initiatives and honorary roles.

Furthermore, the results have implications for police investigations, as they reveal that privacy infringements resulting from data evaluations are determined by the type of data rather than the objective of the evaluation. Notably, as the level of system automation increases, the intensity of privacy invasion decreases. Consequently, it is crucial to ensure a high degree of acceptance of the automation component in systems designed to support criminal investigations. This becomes especially pertinent when incorporating AI assistance, as users' mental models significantly influence their perceptions. Additionally, the findings emphasize the necessity of conducting ethical, legal, and social evaluations of automated mobile forensics in comparison to existing alternatives. While an AI-based approach may present certain challenges in absolute terms, it may be comparatively less problematic than the alternatives currently available. Similarly, when discussing the gradient of automation, debates often focus on complete AI replacement of tasks and jobs, while the optimal solution may involve autonomous user support (Ausat et al., 2023).

In their public communications, authorities employing mobile forensics should emphasize the advantages for civil security while transparently disclosing the initial suspicion and technology-specific benefits. The critical question remains: Whose data are being analyzed and for what purpose? Notably, the data of suspects are considered less sensitive than that of their immediate contacts. Although the latter group may be connected to the crime and their data could assist in its resolution, the qualitative analysis revealed a greater challenge in gaining acceptance for this scenario compared with the analysis of data from entirely unrelated individuals. Interestingly, even those uninvolved persons may still feel they have nothing to hide (Cho et al., 2010).

Overall, a considerable segment of the population in Germany and Austria maintains a favorable perception of mobile forensics. However, this positive view is contingent upon the implementation of effective safeguards against data misuse and the establishment of robust measures to ensure the integrity and resistance to forgery of digital evidence.

5.3. Future research

Further research is imperative to investigate potential variations among indirect users, which may give rise to non-critical combinations, especially concerning automated evaluation and the assessment of image data. Moreover, as we progress, the importance of privacy-compliant AI training will continue to escalate, potentially fostering a novel research domain centered on ethical and environmentally friendly AI training methodologies (Verdecchia et al., 2023). Furthermore, the level of concern pertaining to the aggregation of diverse datasets remains a subject for future investigation.

Author's note

Prior to starting the procedure, the participants were informed that it is of high importance to understand free opinions and attitudes on mobile forensics from the citizens' perspective and that we were very happy if they would share their opinions with us. Still, however we stressed that they are free in taking part or not. Participation in the interview study was completely voluntary, participation in the survey study was reimbursed. Further, we ensured a high standard privacy protection in both studies and let the participants know that none of their answers can be referred to them as persons. Demographic data were also submitted voluntarily and all participants were informed that on request their personal data would be deleted from our encrypted hard drives. After these careful explanations participants reported to feel well informed about the purpose and the aim of the study and their freedom to quit participation at any time. Regarding the privacy policy explanations, the participants reported to understand that high standards were applied and deliberately accepted participation. Participant privacy is a key value that our university has committed itself to uphold. From the comments in the open question fields at the end of the survey, we learnt that those participants were interested in the topic and were keen to look at the results, which we assured them to receive.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://osf.io/ybq9u/?view_only=ecea622e09e343d0baca083823d764bf.

Ethics statement

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. Written informed consent from the participants was not required to participate in this study in accordance with the national legislation and the institutional requirements.

Author contributions

JRH was responsible for designing both empirical studies, conducting both studies, analyzing the data, and drafting the manuscript. E-MS was responsible for the literature review, data analysis of the quantitative study, and drafting of the manuscript. MZ was responsible for general advisory and editing the manuscript. ACV was responsible for designing both empirical studies, drafting the manuscript, and general advisory. All authors contributed to the article and approved the submitted version.

Funding

This study was funded by the German Federal Ministry of Research and Education as a part of the SmartIdentifikation research project (Grant Number 13N14764).

Acknowledgments

We would like to express our gratitude to Linda Juskowiak and Maxime Metzler for their assistance and support in conducting this research. In the course of refining this manuscript, the authors used ChatGPT 4, a language model developed by OpenAI, to enhance the readability of individual paragraphs. The following prompt was provided: You are a copy-editor for an academic journal and provide improved versions of paragraphs in a document. The preferred style is a classic style. In the process, care was taken to ensure that only non-sensitive data were shared with the service. Furthermore, the authors ensured that the revisions made by ChatGPT 4 did not introduce new concepts or any form of novel intellectual property. The authors conscientiously reviewed and edited the output as necessary, and therefore, accept full responsibility for the final content presented in this publication.

Conflict of interest

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

Publisher's note

References

Acquisti, A., Brandimarte, L., and Loewenstein, G. (2015). Privacy and human behaviour in the age of information. *Science* 347, 509–514. doi: 10.1126/science.aaa1465

Al-Dhaqm, A., Abd Razak, S., Ikuesan, R. A., Kebande, V. R., and Siddique, K. (2020). A review of mobile forensic investigation process models. *IEEE Access* 8, 173359–173375. doi: 10.1109/ACCESS.2020.3014615

Alenezi, A. M. (2023). Digital forensics in the age of smart environments: a survey of recent advancements and challenges. *arXiv preprint arXiv:2305.09682*. doi: 10.48550/arXiv.2305.09682

Anandaraj, S., and Kemal, M. (2017). "Research opportunities and challenges of security concerns associated with big data in cloud computing," in *2017 International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC)* (IEEE), 746–751.

Ausat, A. M. A., Massang, B., Efendi, M., Nofirman, N., and Riady, Y. (2023). Can chat GPT replace the role of the teacher in the classroom: a fundamental analysis. *J. Educ.* 5, 16100–16106. doi: 10.31004/joe.v5i4.2745

Baier, D. B., Baier, D., and Brusch, M. (2009). *Conjointanalyse*. Heidelberg: Springer. doi: 10.1007/978-3-662-63364-9

Bansal, G., Zahedi, F. M., and Gefen, D. (2010). The impact of personal dispositions on information sensitivity, privacy concern and trust in disclosing health information online. *Decis. Support Syst.* 49, 138–150. doi: 10.1016/j.dss.2010.01.010

Barmpatsalou, K., Cruz, T., Monteiro, E., and Simoes, P. (2018). Mobile forensic data analysis: suspicious pattern detection in mobile evidence. *IEEE Access* 6, 59705–59727. doi: 10.1109/ACCESS.2018.2875068

Bryant, R., Katz, R. H., Lazowska, E. D. (2008). Big-data computing: creating revolutionary breakthroughs in commerce, science and society. Available online at: https://www.immagic.com/eLibrary/ARCHIVES/GENERAL/CRA_US/C081222B.pdf

Calero Valdez, A., and Ziefle, M. (2019). The users' perspective on the privacy-utility trade-offs in health recommender systems. *Int. J. Hum. Comput. Stud.* 121, 108–121. doi: 10.1016/j.ijhcs.2018.04.003

Carrier, B. (2003). Defining digital forensic examination and analysis tools using abstraction layers. *Int. J. Digit. Evid.* 1, 1–12.

Cho, H., Lee, J.-S. S., and Chung, S. (2010). Optimistic bias about online privacy risks: testing the moderating effects of perceived controllability and prior experience. *Comput. Hum. Behav.* 26, 987–995. doi: 10.1016/j.chb.2010.02.012

Comte, E., and Lavenex, S. (2022). Differentiation and de-differentiation in eu border controls, asylum and police cooperation. *Int. Spectator* 57, 124–141. doi: 10.1080/03932729.2022.2021011

Cruz-Cunha, M. M., and Mateus-Coelho, N. R. (2020). *Handbook of Research on Cyber Crime and Information Privacy*. Pennsylvania:IGI Global. doi: 10.4018/978-1-7998-5728-0

Dhirani, L. L., Mukhtiar, N., Chowdhry, B. S., and Newe, T. (2023). Ethical dilemmas and privacy issues in emerging technologies: a review. *Sensors* 23, 1151. doi: 10.3390/s23031151

Dinev, T., Bellotto, M., Hart, P., Russo, V., Serra, I., and Colautti, C. (2006). Privacy calculus model in e-commerce - A study of Italy and the United States. *Eur. J. Inform. Syst.* 15, 389–402. doi: 10.1057/palgrave.ejis.3000590

Dogan, S., and Akbal, E. (2017). "Analysis of mobile phones in digital forensics," in *2017 40th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)* (IEEE), 1241–1244. doi: 10.23919/MIPRO.2017.7973613

Dritsas, S., Gritzalis, D., and Lambrinoudakis, C. (2006). Protecting privacy and anonymity in pervasive computing: trends and perspectives. *Telemat. Inform.* 23, 196–210. doi: 10.1016/j.tele.2005.07.005

Du, X., Hargreaves, C., Sheppard, J., Anda, F., Sayakkara, A., Le-Khac, N.-A., et al. (2020). "SOK: exploring the state of the art and the future potential of artificial intelligence in digital forensic investigation," in *Proceedings of the 15th International Conference on Availability, Reliability and Security* (New York, NY: Association for Computing Machinery), 1–10. doi: 10.1145/3407023.3407068

Ermakova, T., Baumann, A., Fabian, B., and Krasnova, H. (2014). "Privacy policies and users' trust: does readability matter?" in *AMCIS*. Available online at: https://boris.unibe.ch/68895/

Eurostat. (2022). How many citizens had basic digital skills in 2021? — ec.europa.eu. Available online at: https://ec.europa.eu/eurostat/web/products-eurostat-news/-/ddn-20220330-1 (accessed June 30, 2023).

Federal Ministry of Education and Research (BMBF) (2018). Available online at: https://www.sifo.de/sifo/en/research-projects/society/migration-issues/approved-projects-in-the-field-ivil-security-migration-issues.html (accessed June 30, 2023).

Gantz, J., and Reinsel, D. (2012). The digital universe in 2020: big data, bigger digital shadows, and biggest growth in the far east. *IDC iView* 2007, 1–16.

Gudivada, V. N., Baeza-Yates, R., and Raghavan, V. V. (2015). Big data: promises and problems. *Computer* 48, 20–23. doi: 10.1109/MC.2015.62

Hayes, D., Cappa, F., and Le-Khac, N. A. (2020). An effective approach to mobile device management: security and privacy issues associated with mobile applications. *Digit. Bus.* 1, 100001. doi: 10.1016/j.digbus.2020.100001

Huang, Y., Li, Y. J., and Cai, Z. (2023). Security and privacy in metaverse: a comprehensive survey. *Big Data Mining Anal.* 6, 234–247. doi: 10.26599/BDMA.2022.9020047

ISO (2011). *Systems and Software Engineering – Systems and Software Quality Requirements and Evaluation (SQuaRE) – System and Software Quality Models*. Standard, International Organization for Standardization, Geneva.

Jasserand, C. (2018). Subsequent use of GDPR data for a law enforcement purpose: the forgotten principle purpose limitation. *Eur. Data Prot. Rev.* 4, 152. doi: 10.21552/edpl/2018/2/6

Jindal, D., Kaushik, M., and Bahl, B. (2023). "Emerging trends of privacy and security in cloud computing," in *AIP Conference Proceedings* (AIP Publishing). doi: 10.1063/5.0148999

Johnson, R., and Orme, B. (2003). "Getting the most from CBC," in *Sequim: Sawtooth Software Research Paper Series, Sawtooth Software*.

Karamanidou, L., Kasparek, B., and Hess, S. (2020). Border management and migration control–comparative report. Zenodo. doi: 10.5281/zenodo.3732864

Kimmelmann, N., Miesera, S., Moser, D., and Pool Maag, S. (2022). "Inclusion for all in vet? a comparative overview of policies and state of research about migration, integration and inclusion in Germany, Austria and Switzerland," in *Migration and Inclusion in Work Life–The Role of VET: Emerging Issues in Research on Vocational Eduction & Training*, 117–165.

Krasnova, H., and Veltri, N. F. (2010). "Privacy Calculus on Social Networking Sites: Explorative Evidence from Germany and USA," in *Proceedings of the Annual Hawaii International Conference on System Sciences*, 1–10. doi: 10.1109/HICSS.2010.307

Krippendorff, K. (2018). *Content Analysis: An Introduction to its Methodology*. Sage Publications. doi: 10.4135/9781071878781

Kuckartz, U. (2014). *Qualitative Text Analysis: A Guide to Methods, Practice and Using Software*. Sage. doi: 10.4135/9781446288719

Li, Y. (2011). Empirical studies on online information privacy concerns: literature review and an integrative framework. *Commun. Assoc. Inform. Syst.* 28, 453–496. doi: 10.17705/1CAIS.02828

Losavio, M. M., Chow, K., Koltay, A., and James, J. (2018). The internet of things and the smart city: legal challenges with digital forensics, privacy, and security. *Sec. Privacy* 1, e23. doi: 10.1002/spy2.23

Malheiros, M., Preibusch, S., and Sasse, M. A. (2013). ""Fairly truthful": the impact of perceived effort, fairness, relevance, and sensitivity on personal data disclosure," in *Trust and Trustworthy Computing: 6th International Conference, TRUST 2013* (Berlin; Heidelberg: Springer Verlag), 250–266.

Markos, E., Milne, G. R., and Peltier, J. W. (2017). Information sensitivity and willingness to provide continua: a comparative privacy study of the United States and Brazil. *J. Public Policy Market.* 36, 79–96. doi: 10.1509/jppm.15.159

Marturana, F., Me, G., Berte, R., and Tacconi, S. (2011). "A quantitative approach to triaging in mobile forensics," in *2011IEEE 10th International Conference on Trust, Security and Privacy in Computing and Communications* (IEEE), 582–588. doi: 10.1109/TrustCom.2011.75

Metag, J., and Marcinkowski, F. (2014). Technophobia towards emerging technologies? A comparative analysis of the media coverage of nanotechnology in Austria, Switzerland and Germany. *Journalism* 15, 463–481. doi: 10.1177/1464884913491045

Milne, G. R., Pettinico, G., Hajjat, F. M., and Markos, E. (2017). Information sensitivity typology: mapping the degree and type of risk consumers perceive in personal data sharing. *J. Cons. Aff.* 51, 133–161. doi: 10.1111/joca.12111

Mothersbaugh, D. L., Foxx, W. K., Beatty, S. E., and Wang, S. (2012). Disclosure antecedents in an online service context: the role of sensitivity of information. *J. Serv. Res.* 15, 76–98. doi: 10.1177/1094670511424924

Neyer, F. J., Felber, J., and Gebhardt, C. (2012). Entwicklung und validierung einer kurzskala zur erfassung von technikbereitschaft. *Diagnostica* 58, 87–99. doi: 10.1026/0012-1924/a000067

Nissenbaum, H. (2010). *Privacy in Context: Technology, Policy, and the Integrity of Social Life*. Stanford, CA: Stanford University Press.

Pfisterer, V. M. (2019). The right to privacy–a fundamental right in search of its identity: uncovering the Cjeu's flawed concept of the right to privacy. *German Law J.* 20, 722–733. doi: 10.1017/glj.2019.57

Ribble, M. S., Bailey, G. D., and Ross, T. W. (2004). Digital citizenship: addressing appropriate technology behavior. *Learn. Lead. Technol.* 32, 6.

Rohm, A. J., and Milne, G. R. (2004). Just what the doctor ordered: the role of information sensitivity and trust in reducing medical information privacy concern. *J. Bus. Res.* 57, 1000–1011. doi: 10.1016/S0148-2963(02)00345-4

Rouvroy, A., and Poullet, Y. (2009). "The right to informational self-determination and the value of self-development: reassessing the importance of privacy for democracy," in *Reinventing Data Protection?* (Springer), 45–76. doi: 10.1007/978-1-4020-9498-9_2

Saranya, S., and Usha, G. (2023). "Forensic analysis of online social network data in crime scene investigation," in *Artificial Intelligence and Blockchain in Digital Forensics* (River Publishers), 183–209.

Sawtooth Software, Inc. (2017). "The CBC system for choice-based conjoint analysis," in *Sawtooth Software Technical Paper Series*, 98382. Available online at: https://sawtoothsoftware.com/resources/technical-papers/cbc-technical-paper, (accessed June 30, 2023).

Schomakers, E.-M., Lidynia, C., Müllmann, D., Matzutt, R., Wehrle, K., Ziefle, M., et al. (2021a). Insights on data sensitivity from the technical, legal and the users' perspectives-practical suggestions on how to raise more awareness for the assumed exercise of informational self-determination. *Comput. Law Rev. Int.* 22, 8–15. doi: 10.9785/cri-2021-220103

Schomakers, E.-M., Lidynia, C., Müllmann, D., and Ziefle, M. (2019). Internet users' perceptions of information sensitivity–insights from germany. *Int. J. Inform. Manage.* 46, 142–150. doi: 10.1016/j.ijinfomgt.2018.11.018

Schomakers, E.-M., Lidynia, C., and Ziefle, M. (2020). All of me? Users' preferences for privacy-preserving data markets and the importance of anonymity. *Electron. Mark.* 30, 649–665. doi: 10.1007/s12525-020-00404-9

Schomakers, E.-M., Lidynia, C., and Ziefle, M. (2021b). The role of privacy in the acceptance of smart technologies: applying the privacy calculus to technology acceptance. *Int. J. Hum. Comput. Interact.* 38, 1276–1289. doi: 10.1080/10447318.2021.1994211

Schomakers, E.-M., and Ziefle, M. (2019). "Privacy concerns and the acceptance of technologies for aging in place," in *International Conference on Human-Computer Interaction* (Orlando, FL: Springer), 313–331.

Selting, M., and Auer, P. (2011). A system for transcribing talk-in-interaction: Gat 2 translated and adapted for english by elizabeth couper-kuhlen and dagmar barth-weingarten. *Gesprächsforschung–Online-Zeitschrift zur verbalen Interaktion* 12, 1–51.

Sikos, L. F. (2021). AI in digital forensics: ontology engineering for cybercrime investigations. *Wiley Interdiscipl. Rev. Forensic Sci.* 3, e1394. doi: 10.1002/wfs2.1394

Simsa, R. (2017). Leaving emergency management in the refugee crisis to civil society? The case of Austria. *J. Appl. Sec. Res.* 12, 78–95. doi: 10.1080/19361610.2017.1228026

Staiano, J., Oliver, N., Lepri, B., de Oliveira, R., Caraviello, M., and Sebe, N. (2014). "Money walks: a human-centric study on the economics of personal mobile data," in *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (New York, NY), 583–594.

Tene, O., and Polonetsky, J. (2012). Big data for all: privacy and user control in the age of analytics. *Nw. J. Tech. Intell. Prop.* 11, 27.

Verdecchia, R., Sallou, J., and Cruz, L. (2023). A systematic review of Green <scp>AI</scp>. *WIREs Data Mining and Knowledge Discovery*. doi: 10.1002/widm.1507

Voigt, P., and Von dem Bussche, A. (2017). "The EU general data protection regulation (GDPR)," in *A Practical Guide, 1st edn.* (Cham: Springer International Publishing), 10–5555.

Wilkowska, W., Offermann-van Heek, J., Florez-Revuelta, F., and Ziefle, M. (2021). Video cameras for lifelogging at home: preferred visualization modes, acceptance, and privacy perceptions among German and Turkish participants. *Int. J. Hum. Comput. Interact.* 37, 1436–1454. doi: 10.1080/10447318.2021.1888487

Xu, H., Dinev, T., Smith, H. J., and Hart, P. (2008). "Examining the formation of individual's privacy concerns: toward an integrative view," in *International Conference on Information Systems* (Paris), 6.

Ziefle, M., Halbey, J., and Kowalewski, S. (2016). "Users' willingness to share data on the internet: perceived benefits and caveats," in *IoTBD* (Stuttgart: Science and and Technology Publications (SCITEPRESS)), 255–265. doi: 10.5220/0005897402550265

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 12 October 2023
DOI 10.3389/fpsyg.2023.1017313

OPEN ACCESS

EDITED BY
Andrea Szymkowiak,
Abertay University, United Kingdom

REVIEWED BY
Shilpa Kumari,
Banaras Hindu University, India
Hitesh Mohan,
Lovely Professional University, India
Shi Yin,
Hebei Agricultural University, China

*CORRESPONDENCE
John Protzko
✉ protzko@gmail.com

RECEIVED 11 August 2022
ACCEPTED 20 June 2023
PUBLISHED 12 October 2023

CITATION
Protzko J and Schooler JW (2023) What I didn't grow up with is dangerous: personal experience with a new technology or societal change reduces the belief that it corrupts youth.
Front. Psychol. 14:1017313.
doi: 10.3389/fpsyg.2023.1017313

What I didn't grow up with is dangerous: personal experience with a new technology or societal change reduces the belief that it corrupts youth

John Protzko[1,2]* and Jonathan W. Schooler[2]

[1]Department of Psychological Science, Central Connecticut State University, New Britain, CT, United States, [2]Department of Psychological & Brain Sciences, University of California, Santa Barbara, Santa Barbara, CA, United States

Introduction: Throughout history, technological and societal changes consistently receive suspicion. Their influences appear damaging, corrupting, and potential precursors to societal downfall, with today's youth often portrayed as the primary victims. This study aims to explore an underlying reason for these perceptions and to investigate why society frequently perceives technological and societal transitions as detrimental to the younger generation.

Methods: We conduct two studies across a total of 1,702 participants. In a pilot study, American adults generate a list of technological/societal innovations they believe to be especially problematic for youth in various ways. The second study maps beliefs that specific technological/societal shifts are corruptive, correlating with whether American adults experience them during their upbringing.

Results: People view recent technologies as particularly corrupting of today's youth. A notable within-person correlation exists between an individual's exposure to specific technologies during their youth and their belief that these technologies corrupt today's youth. Specifically, people are more inclined to view technological/societal shifts as corruptive if they don't experience them during their formative years ($b = -0.09$, $p < 0.001$, 95%CI = [−0.11, −0.09]). When reminded of their own exposure to a particular innovation during their upbringing, however, this relationship reduces.

Discussion: These findings suggest unfamiliarity currently stands as a pivotal factor in societal apprehensions regarding new technological and societal evolutions. As society welcomes new innovations, an enduring cycle emerges where those unacquainted changes seem corruptive to the newer generations. Recognizing this bias, primarily driven by mere unfamiliarity, may be crucial for more balanced evaluations of the inevitable technological and societal progress.

KEYWORDS

societal decline, digital-technology use, social media, screen time, preregistered, moral panic, generational perceptions, cultural adaptation resistance

Introduction

People have been complaining about the corruption of the youth and the decline of society for thousands of years (Smart, 1836; Freeman, 1907). These apparent declines have, throughout history, been blamed on changes in society and technology. Social Media in the early 21st century has been seen as corrupting the youth (e.g., Marwick, 2008; Twenge, 2017), yet concerns about the devastation

the radio has on children (e.g., Preston, 1941; Wartella and Jennings, 2000) and the panic about radio's ability to create a single mass culture with no dissenting opinions (e.g., Davis, 1965; Swingewood, 1977) have faded. Nevertheless, in the early 20th century radio was seen as an invading force that crowded out more intellectual past times such as reading (Eisenberg, 1936; Wartella and Jennings, 2000). Reading, however, was the enemy in the 18th and 19th centuries, leading to the belief that novels led to a frittering away of the young mind (e.g., Hitchcock, 1710; see also Furedi, 2015; Proulx, 2019), while people in the early 21st century complain that the youth do not read enough (Protzko and Schooler, 2019). It appears that we cannot keep our stories straight across generations about *which* technological/societal changes are actually corruptive of the youth. Why do we keep accusing the technologies and societal changes of the day of corrupting the youth?

Shifting foci

Examples of historical complaints about the corrupting influences of then-present technological/societal changes abound and are entertaining to review (see Orben, 2020, for more examples). Socrates criticized writing as a technology that would degrade the youth as it "will create forgetfulness in the learners' souls, because they will not use their memories; they will trust to the external written characters and not remember of themselves" (Plato, 360bce/Jowett, 2019). One would be hard pressed to find people complaining about the corrupting influence of books, the printing press, or writing in the present.

There is the persistent and historically pervasive belief that 'kids these days' are in decline, dating back to at least the 5th Century bce (e.g., Protzko and Schooler, 2019). People largely view the youth of 'the present,' regardless of what present it is, as deficient compared to previous generations (see also Trzesniewski and Donnellan, 2014; Protzko and Schooler, 2022). If the youth of 'the present' are seen as in decline, people may want something to blame.

Historically, explanations for the decline of the youth have tended to focus on contemporary technological/societal changes. The focus often seems to be on 'new' technological/societal changes, not ones that have been around for generations. Complaints about the printing press, corrupting people with bombardments of information (e.g., Gessner, 1565, as cited in Blair, 2003), were largely limited to when the printing press was a newer societal advance. At the time, complaints were no longer about writing in general (as Socrates bemoaned around 1,200 years earlier) but about the new alleged problems introduced when reading was possible *en masse*.

Numerous reports have mocked these historic complaints, and some have even related such historic concerns to modern panics over contemporary technological/societal advances (see Bell, 2010; Gillard, 2018; see also Smiley and Fisher, 2022), yet the warnings of the new technological/societal change persist (e.g., Twenge et al., 2020). Why the constant panic? We propose a lack of personal experience with the technology/societal change as one psychological reason why adults perennially view new technology/societal change as a source of the decline of the youth of the day.

Personal experience

Social Media arose and grew in popularity in the 1990s and 2000s, becoming ubiquitous in the 2010s (Pew Research Center, 2021). People born during the 1980s have the experience of growing up without *and with* such technologies, depending on when they adopted the technology. People born in the 2000s–2010s have little to no experience of not living in an internet and social media-connected world. People born before the 1980s have no personal experience growing up with the internet or social media. This lack of personal experience may color people's understanding of the technology or social change.

The stark differences in the experiences of individuals who grow up with versus without a particular technology raises the following possibility: individuals who lack access to a given technological/societal change while growing up may perceive grave risks of that technology for youth of 'the present' who are exposed to it. This would explain why people in the early 20th century, lacking any experience growing up with the radio, would have seen it as such a corrupting influence on youth (e.g., Eisenberg, 1936; Wartella and Jennings, 2000) while it may not be seen as dangerous to people in the early 21st century, as nearly all Americans have the experience of growing up with the radio.

Personal exposure may help explain why we see the next technology/societal change as dangerous to the youth of 'the present' while being much less concerned with those that we grew up with ourselves. Of course, one's generation and their exposure to a technology is often confounded, as one cannot experience a technology that was not available. Nevertheless, the fact that individuals within a generation can vary in experience with particular technologies enables us to potentially unconfound these two variables. Specifically, the personal exposure hypothesis predicts that within a generation those who were exposed to a particular technology should be less inclined to perceive today's youth as at risk from that particular technology relative to those who had less exposure.

Importantly, we make no claims about the empirical status of *whether* a given technological/societal change is itself harmful (see, Orben and Przybylski, 2019 for effect sizes from correlational work, for example), but rather seek to understand a psychological underpinning of why society's focus keeps shifting across 'corrupting influences'. To better understand how exposure to technologies may impact people's assessments of the challenges facing "kids these days (Protzko and Schooler, 2019, 2022) we investigated two questions: (1) In general, are anecdotal reports accurate in suggesting that adults perceive children at greater risk from newer technologies relative to ones that were available to their generation (i.e., is social media perceived as being more dangerous than television)? (2) Are individuals who themselves lacked experience with a specific technology/societal advance especially inclined to perceive that particular technology as problematic?

As people have been rejecting new technologies and innovations since at least Plato's day (Plato, 360bce/Jowett, 2019) while remaining quieter on the perceived dangers of older technologies, we conduct a systematic study of the possibility that people are more likely to see 'current' technologies as dangerous to the youth than older technologies. Assuming this relationship is shown to be true, we then try to understand one explanation for the distrust of 'current' technologies—people's personal experience with the technology. We propose that a lack of experience with a technology or social change leads people to think it is dangerous for the youth of 'the present', which leads to the following straightforward prediction: for any given advance those who grew up with it will perceive it as being less problematic than those who did not.

Methods

We first ran a pilot study to elicit user-generated technological/societal influences that Americans believe are responsible for corrupting the youth. Participants were asked to identify the reasons or causes of the decline of the youth of 'the present'. Participants were randomly assigned to read one of nine different questions: either a general or a specific question referring to one of eight particular potential forms of decline. Participants reading the general decline question read:

> Children and youths today appear noticeably in decline from the standards of youth of the past. We would like to know what you think the reasons or causes of this decline are.
>
> Below are five open spaces. Think about what you think is causing a decline in the youth of today. For each line, please write less than five words. So for example, if you think 'social media' is one reason, write that below. Please write whatever you honestly believe are the causes or reasons.

For the specific declines, participants were told that children and youths today appear in decline on one of the following domains: they seem to: *be getting more narcissistic*, *be reading less*, *be more politically extreme*, *be lazier*, *lack the desire to work hard*, *be less respectful of authority*, *be too sensitive and politically correct*, *be more violent*. This list was chosen to reflect commonly heard complaints against the youth (e.g., Trzesniewski and Donnellan, 2014; Protzko and Schooler, 2019, 2022) and was not meant to be comprehensive. The specific questions were meant to make the task easier for participants. We collected the responses that all participants gave and reviewed them to try to locate additional, commonly occurring themes. No formal analysis was undertaken.

Participants were 202 members of Amazon Mechanical Turk (mTurk) who were given the survey from November 10, 2019 to November 12, 2019. As there was no confirmatory analysis planned or effect size sought or inferential statistics applied in this pilot study, no power analysis was conducted. See materials and data at https://osf.io/4zk9e/. The responses from this pilot study were used to populate potentially corrupting technologies and societal changes for the main study.

Main study: does not growing up with a technology increase the belief it is corrupting the youth?

The purpose of this study was to relate exposure to certain technological/societal advances in childhood (identified both historically and in the pilot study) to the belief that those advances are corrupting youth of 'the present'.

Methods

Materials

We populated our list of technologies and societal advances from historical complaints, contemporary complaints, and user-identified causes from the pilot study. Participants were given the list twice, once when they were asked what they believe contributes to the decline and corruption of the youth, and again when they identified whether they personally had access to any of these technologies or societal advances growing up. All items were presented in random order. The full list of 'corrupting' technologies and societal advances can be seen in Table 1, along with the base overall rate at which people believed the influence had a corrupting influence.

Procedure

Participants were first be asked what year they were born on a drop down list from 2001 until 1918. Next, participants were randomly assigned to fill out the list of what they had growing up or what they believe is causing a decline in the youth scales in random order. For the *What I Had* scale, participants read:

> Below is a list of different technologies and aspects of society.
>
> Please select below all of the items that you personally had or experienced growing up.
>
> So for example, if TV was around while you grew up but you never watched it, you would not check Television.

For the *Corruption* scale, participants read:

TABLE 1 Univariate percentages of American adults who believe each item is specifically corrupting the youth of 'the present.'

Technology or societal element	% Who believe it is corrupting the youth
Social Media	72.9
Smart phones	53.9
The Internet	58.8
The Radio	5.3
Television	31.0
Reading novels	2.1
Driving cars	4.4
Single parent families	48.6
Video Games in the home	46.9
Heavy Metal music	14.9
24-h news	12.1
Dance clubs	6.5
Nicotine vaporizers	40.5
Netflix	14.1
Jazz music	1.5
Long hair	3.4
Ballroom dancing	1.3
Motion pictures	14.1
Not going to church	37.3
Online dating	22.1
Calculators	2.9
Autocorrect	7.5
Word processors	2.8
Audio/Electronic books	4.3

> Children today appear to be in decline compared to the way children were when you were a child.
>
> We are interested in what you believe contributes to this decline. Below is a list of possible causes. Please select as many as you honestly believe are contributing to the decline of the youth of today.

Participants were given the list of items, presented in random order, and allowed to check as many options as they saw fit.

Analysis plan

The analysis is a within-person analysis, with the prediction that if someone had exposure to a given technological/societal change growing up, they would be less likely to view it specifically as corrupting the youth. This within-subjects mixed effects model includes whether the individual believes the technological/societal change is corrupting as the binary dependent variable, and whether they had exposure to the technology as the independent variable, run with robust standard errors. The model was not able to converge using the technological/societal change as a random-effect, so it was included as a fixed-effect. Analysis scripts, data, and all materials are available at https://osf.io/4zk9e/. This study was preregistered prior to data collection at https://osf.io/yrzxw.

Participants

We collected 1,500 participants, drawn in a stratified way with unequal probabilities of selection, so that the people who complete each survey will resemble the nation's adult population (according to the most recently available Current Population Survey, conducted by the U.S. Census Bureau) in terms of gender, age, education, ethnicity (Hispanic vs. not), race (allowing each respondent to select more than one race), region, and income. The data was collected in December of 2019. The sample size was determined as this study was part of a project running studies with fixed sample sizes at $N = 1{,}500$ (Protzko et al., 2020).

Results

First, as an exploratory analysis, we investigated whether there was a relationship between the year of a technology/social advance's invention and the extent to which it is seen as corrupting. Although not preregistered, this analysis naturally arises from the conjecture that personal experience impacts perceptions of which technology/societal innovations are especially corrupting or benign. We coded each year that each technology/societal advance was invented. In three cases where there was no introduction date (i.e., Long hair; Single-parent homes; Not going to Church) so dates were set to a time seen as generally coinciding with the period in which that societal change became of note: 1969 for Long Hair to coincide with the Hippie movement; 1989 for Single-parent homes to coincide with the popular scapegoating of rising crime in the late 1980s and early 1990s in the U.S. on single-parent homes (e.g., Cohen, 2012); and 2019 for Not going to church to represent the lowest attendance rate at church in the U.S. to that point (Jones, 2021). People in general believed that more modern technologies/societal changes are more corrupting than those that had been invented or popularized earlier ($r_S = 0.67$, $p < 0.001$; see Figure 1). Note that removing the three items that did not have introduction dates does not alter the strength of significance of the relationship ($r_S = 0.65$, $p = 0.001$).

We turn now to our preregistered examination of whether people think a technology or societal advance is more corrupting if they did not experience it themselves growing up. A univariate, item-by-item, analysis could be susceptible to an age confound, whereby older participants are less likely to be exposed to a given technology growing up (e.g., Netflix) and due to their age also happen to hold more negative views of the youth (e.g., Protzko and Schooler, 2022). This confound, however, would not operate on the within-person analysis, as such general negative views (between-person effects) are statistically divorced from the within-person effect of being exposed to a given technology of societal advance growing up. Therefore, a mixed-effects analysis is able to disentangle the within-person effect (context for growing up exposed to a specific technology) from the between-person effects (random intercepts of tendency to believe in overall corrupting influences).

Results from this overall mixed-effect model confirmed our hypothesis. If someone did not have the context of being exposed to a given technology or societal advance growing up, they were more likely to believe that it has a corrupting influence on the youth of 'the present' ($b = -0.09$, $SE_{robust} = 0.01$, $p < 0.001$, 95%CI = [−0.11, −0.09]; see Figure 2). This relationship did not change when including the item as a fixed effect in the model ($b = -0.04$, $SE_{robust} = 0.01$, $p < 0.001$, 95%CI = [−0.05, −0.03]). This corresponds to a 56% increase, across technological/societal changes, in the belief that something corrupts the youth of 'the present' if someone did not experience it themselves growing up.

We also observed that the older someone is, the more likely they were to think that technological/societal changes are corrupting ($b = 0.001$, $SE_{robust} = 0.0002$, $p < 0.001$, 95%CI = [0.0005, 0.001]). Nevertheless, conditioning on age ($M_{age} = 50.1$, range 18–87) did not change the results: people were still more likely to think a technological/societal changes was more corrupting if they did not experience it growing up ($b = -0.04$, $SE_{robust} = 0.01$, $p < 0.001$, 95%CI = [−0.05, −0.03]). Thus, overall, older participants are more likely to see corruption of the youth, but even then having not grown up experiencing a given technological/social change corresponds to an increase in the belief that it is corrupting the youth, on average.

Finally, participants were randomly assigned to fill in what they believe contributes to the corruption of the youth either first or after describing what they had access to as a child. The relationship between personal experience and beliefs of corruption was stronger when participants indicated were first asked what they were exposed to as a child ($b = -0.05$, $SE_{robust} = 0.01$, $p < 0.001$, 95%CI = [−0.07, −0.04]) compared to when they were first asked why they believe corrupts the youth first ($b = -0.03$, $SE_{robust} = 0.01$, $p < 0.001$, 95%CI = [−0.04, −0.01]). Thus, the simple act of reminding someone that they had access to a given technological/societal advance when they were young is enough to reduce the belief that that particular advance is corrupting the youth of the present.

As an exploratory analysis, we included data on the highest level of school participants had completed or the highest degree they had received. Response options were No High School Diploma (less than 12th grade), High School Graduate- high school DIPLOMA or the equivalent (For example: GED), Associate degree in college-Occupational/vocational program or Academic program, Bachelor's degree (For example: BA, AB, BS), Master's or Professional School degree (For example: MA, MS, MEng, MEd, MSW, MBA), Doctorate Degree (For example: PhD, EdD, MD, DDS, DVM, JD). We found that those who had a higher education level were slightly less likely to

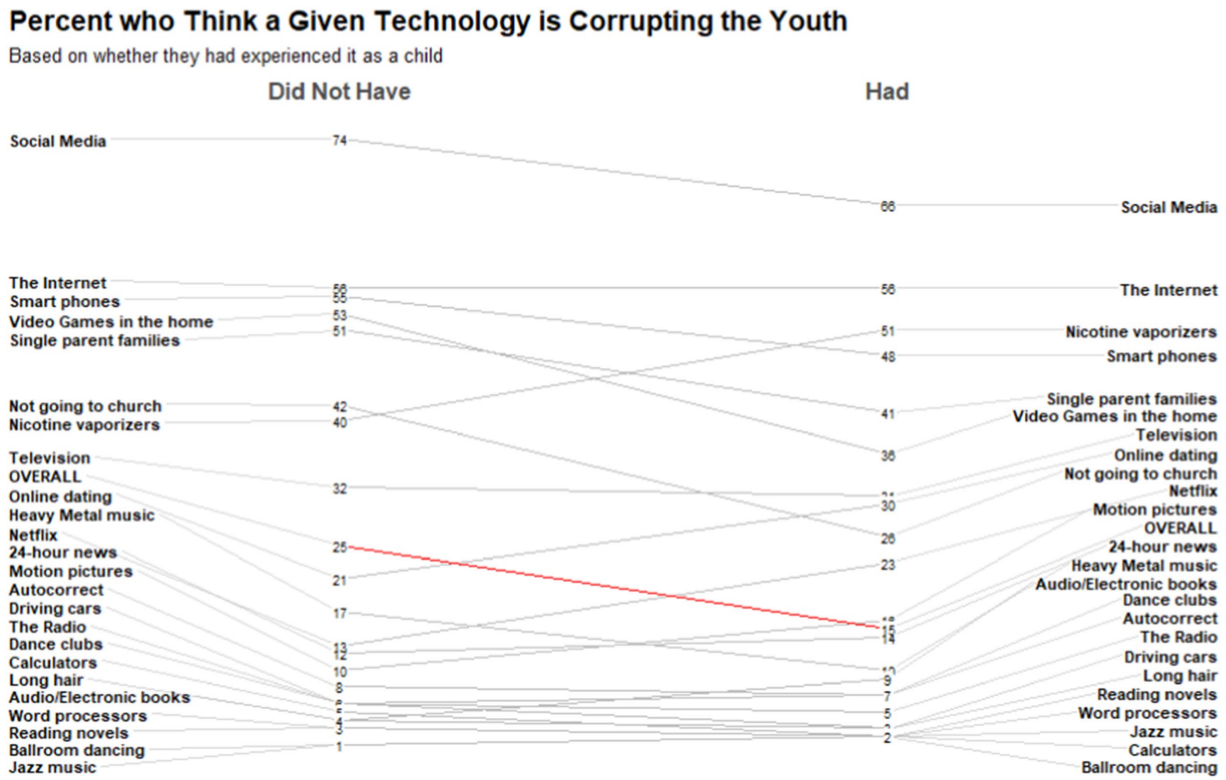


FIGURE 1
Year of Invention (or Popularization, denoted by an asterisk) and the extent to which people in general think those technology/societal advances are corrupting the youth. Line is a lowess line; gray bars are 95%CI that start once enough data becomes available.

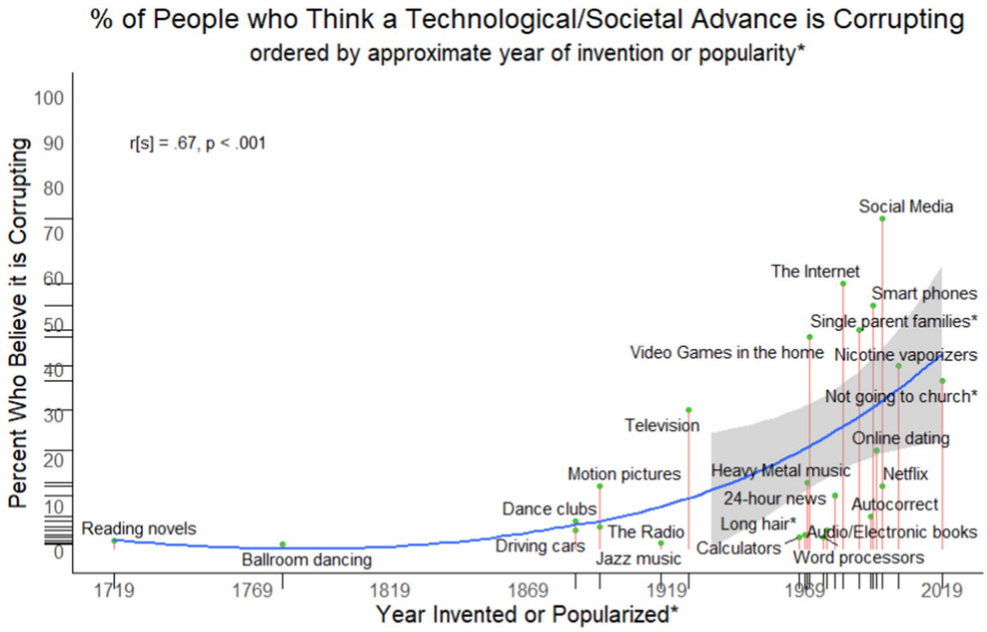


FIGURE 2
Slope graph of the percent of people who believe that a given technology or societal advance is corrupting the youth of 'the present', based on whether they had personal exposure to the technology or societal advance in their own childhoods. Red line is the overall model effect, people were 56% more likely to believe a given technology/societal advance was corrupting, on average, if they had not personally experienced it growing up.

believe that technology was corrupting ($b = -0.005$, $SE_{robust} = 0.001$, $p < 0.001$, 95%CI = [−0.01, −0.003]). Including education as a covariate did not alter the results that those who grew up experiencing a technology found them to be less corrupting ($b = -0.1$, $SE_{robust} = 0.01$, $p < 0.001$, 95%CI = [−0.11, −0.09]). Indeed, even restricting the dataset to just those with the highest level of education (doctoral degrees, $n = 63$) still shows that those who experienced a given technology/societal advance growing up were less likely to believe it was corrupting ($b = -0.11$, $SE_{robust} = 0.02$, $p < 0.001$, 95%CI = [−0.16, −0.07]). Thus, while those who are of higher education may be less likely to believe technologies/societal advances are corrupting in general, they are just as susceptible to believing what they did not grow up with is dangerous.

Discussion

We first confirmed that technological/societal changes that are more recent are seen as more dangerous to the youth than those of the past. So, for example, social media, the internet, and smartphones were seen as corrupting the youth of the present, while driving cars, jazz music, and dance clubs are no longer seen as dangerous. Secondly, we demonstrated that people generally believe that technologies and societal innovations they did not personally experience growing up are uniquely dangerous to today's youth. Overall, participants who did not grow up experiencing a particular technological/social change showed a 56% increase in believing that specific aspect of life is corrupting the youth of 'the present' relative to participants who did grow up with it.

These findings suggest that newer technologies/societal advances are often the targets of concern for the very reason that we lack the personal experience of growing up with them. American adults held varying beliefs on the corrupting influence of the items we presented. Social Media was seen as the most corrupting influence (74%) on the youth of 'the present', for example, but if someone grew up experiencing social media, they found it a considerably less dangerous force (66%). The same is also said about growing up being exposed to Heavy Metal music; people who grew up not experiencing Heavy Metal music see it as particularly corrupting (17%), but much less so when they experienced heavy metal themselves (10%).

As shown in Figure 2, technologies/social changes that were invented or popularized more recently are more likely to be seen as dangerous. Our list was created from lists of historical complaints, as well as generated freely by participants. Therefore, at some point in time people believed these items to be dangerous. We confirmed that a number of technological/social changes that have historically been seen as corrupting or dangerous to the youth when they first appeared are no longer thought to be dangerous. In 1843, for example, warnings were raised about the dangers of ballroom dancing to the youth: "If you wish to preserve in its freshness their modest innocence…suffer them not to waltz" (The Waltz, 1843, p. 152); yet in 2019, only 1% of American adults saw ballroom dancing as a corrupting influence on the youth of 'the present'. Thus, there is considerable heterogeneity across what people believe is corrupting, both presently and historically.

Why do we keep believing each new technological/social change comes with it a danger to the youth? Here we start to explore one reason, lack of personal experience. Someone growing up in a single parent house has context for what it is like, someone who grew up in a dual parent house does not, for example. Presumably, when someone has personal experience with a technology/societal advance they may see it is as less problematic than they might have otherwise thought. They also may be more reluctant to express concern about growing up with a technology/innovation that might bear on their own proficiencies.

As people get older, they are more critical of the youth of 'the present' (Protzko and Schooler, 2022). We also show here that as people age, they tend to see more technological/societal changes as corrupting the youth overall. This may be because as people age their memories for their childhood become more favorable (Field, 1997) and they view the past as more idyllic (Eibach et al., 2003; Protzko and Schooler, 2019; see also Mastroianni and Gilbert, 2023). Thus, not only do youth of 'the present' seem in decline as people age, but society seems in decline; older people are more likely to search for explanations involving technologies/societal changes they do not have context growing up with.

People are similarly critical of technology that specifically was invented after they were born (Smiley and Fisher, 2022). Our work here connects to this literature and helps explain why. Technology invented while one is alive is potentially less likely to have been experienced directly while growing up, especially if that technology was invented in one's adulthood. This may be why so many older adults simultaneously have social media accounts yet believe those social media sites are dangerous for children. A lack of context for social media while growing up makes the societal advance seem dangerous.

We also found an important mitigation strategy, simply reminding people of what technologies/societal advances they grew up with is associated with a decreased belief in the corrupting influence, strengthening the bond between experience and belief. This presents the intriguing possibility that simply being reminded of what one was actually exposed to as a child may reduce the belief that those technologies or societal changes are corrupting. Similar research has explored the idea that such simple reminders can reduce beliefs in the decline of society; asking people to reflect on how their own driving ability has changed in the past 10 years reduces the belief that other people's driving ability has become more aggressive, reckless, for example (see Eibach and Libby, 2009; see also our replications at https://osf.io/xrbfp/). Our results may extend future work about the power of simple contextual reminders on reducing prejudice against the youth and panics about the next technological/societal change.

Implications

Of course, the fact that personal experience with a societal innovation reduces people's concern about its impact should not necessarily allay fears about any particular new technology or societal advance. Similarly, just because people are generally less uneasy about older innovations relative to newer ones does rule out the possible emergence of a truly dangerous new innovation. Nevertheless, fostering a general awareness of the present findings may help to contextualize the concern that current technologies are harming today's youth. Remembering that adults were similarly distressed about the technologies that we used as youth, and understanding that part of the reason we see new technologies and societal changes as corrupting of the youth of 'the present' is that we personally lacked exposure to them growing up, could temper our fears. More generally, in debates, both present and future, about the dangers of new technologies and societal advances, the work here can be invoked to argue that the worries and intuitive sense of danger we will continue to fear

are not necessarily accurate. Those fears may be influenced by a simple lack of exposure to the new advance.

Concerns about the corrupting impact of new technologies might also be dampened by reducing people's ill-founded distress about the youth of the day. As noted Protzko and Schooler (2019) demonstrated that people routinely denigrate the youth of the day for reasons unrelated to the characteristics of the population at large, but rather due to qualities of themselves. Essentially people assume the youth of today are lacking in whatever particular qualities they themselves excel. Reducing our tendency to view the youth of the present as deficient may help to alleviate some of our concerns about the negative impact of new technology or social change.

Limitations and future directions

The present findings are wholly consistent with the view that whether one grew up with a particular technology/societal change impacts their belief in its corruptive impact of the youth of 'the present'. It must be acknowledged, however, that this observed relationship was largely correlational. It is not random who takes up a new technology or societal advance. People who are more interested in the technology or who are more receptive to the societal change are undoubtedly more likely to be early adopters. Nevertheless, we did find an effect of whether or not people were first asked to consider their own personal experience with an innovation on their assessment of its impact. The mitigating effect of remembering that one personally used a technology when growing up further helps to build a causal case that a lack of personal experience with an innovation contributes to its perceived danger.

Future work can look for exogeneous exposure of new technologies (e.g., rollouts in certain markets but not others) to further test the causal implications of these findings. It should be the case that if one is randomly exposed to a new technology growing up (because one is in a test area where the technology is available) that this will lead to less concern about it relative to those who were not exposed to it. Furthermore, future work can explore the causal effect of presenting the findings here on tempering fears. Experiments could test whether alerting people to the relationship between increased exposure and reduced fear can cause people to be less concerned about the new technology or social change under debate.

There may also be a concern about the prompt used in the main study where participatns were told: "Children and youths today appear noticeably in decline from the standards of youth of the past. We would like to know what you think the reasons or causes of this decline are." before being asked what they believed the cause of any declines are. We believe this is not an issue of a demand effect for a number of reasons. First, demand effects have been shown to not reliably exist in research conducted with non-student samples (for a meta-analysis, see: Coles and Frank, 2023). Second, the belief that the youth are in decline is pervasive and has been demonstrated in numerous studies (e.g., Trzesniewski and Donnellan, 2014; Protzko and Schooler, 2019, 2022), and prompting would likely not induce such a belief. Third and most importantly, this would not alter the within-person relationship between not experiencing a given technology and thinking that technology is particularly dangerous. Still, future research, especially if using student samples, should be aware of such possible demand effects.

Conclusion

We can never know what it is like to grow up any other way than we did. We only have our own experience, no more and no less. This experience apparently matters for later beliefs about society. We are more likely to see a given technology or societal change as a corrupting, damaging force on the youth of 'the present' if we did not have the context for what it was like to experience it growing up. As society changes more rapidly and technological innovations become more frequent (e.g., Kurzweil, 2004), we will continue to find ourselves in a world that looks different from the one we grew up in—seeing danger for the next youth. This apparent recurring process—of innovations being spurned as corrupting by older generations who did not grow up with them—could continue in perpetuity. Considering one's own experience with a technology, however, serves to mitigate projections of its negative impact on the young, potentially suggests one way to escape this perennial cycle. If older generations remember that they thrived in the context of the novel technologies and societal advances of their day, they may gain a more optimistic vision that current youth can similarly prosper despite or perhaps even because of the new developments that they grow up with.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found at: https://osf.io/4zk9e/.

Ethics statement

This study was found exempt by the University of California, Santa Barbara Office of Research and Human Subjects (IRB).

Author contributions

JP came up with the idea, programmed the study, and collected and analyzed the data. All authors refined the methods, wrote the original draft, and revised the manuscript.

Funding

The research was funded by the John E. Fetzer Memorial Trust.

Acknowledgments

We thank the reviewer for their suggestions to improve the article.

Conflict of interest

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

Publisher's note

References

Bell, V. (2010). Don't Touch That Dial! A history of media technology scares, from the printing press to Facebook. Slate. Available at: https://slate.com/technology/2010/02/a-history-of media-technology-scares-from-the-printing-press-to-facebook.html

Blair, A. (2003). Reading strategies for coping with information overload ca. 1550–1700. *J. Hist. Ideas* 64, 11–28. doi: 10.1353/jhi.2003.0014

Cohen, P. (2012). Single moms can't be scapegoated for the murder rate anymore. The Atlantic. Available at: https://www.theatlantic.com/sexes/archive/2012/11/single-moms-cant-be-scapegoated-for-the-murder-rate-anymore/265576/

Coles, N. A., and Frank, M. C. (2023). A quantitative review of demand characteristics and their underlying mechanisms. Available at: https://psyarxiv.com/uw85a

Davis, R. E. (1965). *Response to innovation: a study of popular argument about new mass media*. Iowa City, United States: The University of Iowa.

Eibach, R. P., and Libby, L. K. (2009). CHAPTER 16 ideology of the good old days: exaggerated perceptions of moral decline and conservative politics. *Soc. Psychol. Bases Ideol. Syst. Justif.*, 402–424. doi: 10.1093/acprof:oso/9780195320916.003.016

Eibach, R. P., Libby, L. K., and Gilovich, T. D. (2003). When change in the self is mistaken for change in the world. *J. Pers. Soc. Psychol.* 84, 917–931. doi: 10.1037/0022-3514.84.5.917

Eisenberg, A. L. (1936). *Children and radio programs: A study of more than three thousand children in the New York metropolitan area*. New York, NY Columbia University Press.

Field, D. (1997). Looking back, what period of your life brought you the most satisfaction. *Int. J. Aging Hum. Dev.* 45, 169–194.

Freeman, K. J. (1907). *Schools of hellas: An essay on the practice and theory of ancient Greek ducation from 600 to 300 BC*. London, England: Macmillan.

Furedi, F. (2015). The media's first moral panic. *History Today* 65, 46–48.

Gillard, J. (2018). The 2,500-year-old history of adults blaming the younger generation. History Today. Available at: https://historyhustle.com/2500-years-of-people-complaining-about-the-younger-generation/

Hitchcock, E. (1710). *Memoirs of the bloomsgrove family. In a series of letters to a respectable citizen of Philadelphia. Containing sentiments on a mode of domestic education, suited to the present state of society, government, and manners, in the United States of America: And on the dignity and importance of the female character.: Interspersed with a variety of interesting anecdotes*. Printed at Boston,:: by Thomas and Andrews, at Faust's Statue, no. 45, Newbury Street.

Jones, J. M. (2021). U.S. church membership falls below majority for first time. Gallup. Available at: https://news.gallup.com/poll/341963/church-membership-falls-below-majority-first-time.aspx

Kurzweil, R. (2004). "The law of accelerating returns" in *Alan Turing: Life and legacy of a great thinker* (Berlin, Heidelberg: Springer), 381–416.

Marwick, A. E. (2008). To catch a predator? The MySpace moral panic. First Monday.

Mastroianni, A. M., and Gilbert, D. T. (2023). The illusion of moral decline. *Nature*, 618, 782–789. doi: 10.1038/s41586-023-06137-x

Orben, A. (2020). The Sisyphean cycle of technology panics. *Perspect. Psychol. Sci.* 15, 1143–1157. doi: 10.1177/1745691620919372

Orben, A., and Przybylski, A. K. (2019). The association between adolescent well-being and digital technology use. *Nat. Hum. Behav.* 3, 173–182. doi: 10.1038/s41562-018-0506-1

Pew Research Center (2021). Internet/Broadband Fact Sheet. Available at: https://www.pewresearch.org/internet/fact-sheet/internet-broadband/

Plato (360bce). Phaedrus. Trans. Benjamin Jowett. Oxford, 1892. Project Gutenberg. Ed. Michael Hart. Available at: http://www.gutenberg.org/ebooks/1636

Preston, M. I. (1941). Children's reactions to movie horrors and radio crime. *J. Pediatr.* 19, 147–168. doi: 10.1016/S0022-3476(41)80059-6

Protzko, J., and Schooler, J. W. (2019). Kids these days: why the youth of today seem lacking. *Sci. Adv.* 5:eaav5916. doi: 10.1126/sciadv.aav5916

Protzko, J., and Schooler, J. W. (2022). Who denigrates today's youth?: the role of age, implicit theories, and sharing the same negative trait. *Front. Physiol.* 13:723515.

Protzko, J., Krosnick, J., Nelson, L. D., Nosek, B. A., Axt, J., Berent, M., et al. (2020). High replicability of newly-discovered social-behavioral findings is achievable. Available at: https://psyarxiv.com/n2a9x/download?format=pdf

Proulx, F. (2019). *Victims of the book: reading and masculinity in fin-de-siècle France*. Toronto, Canada: University of Toronto Press.

Smart, C. (1836). *The works of Horace, translated literally into English prose; for the use of those who are desirous of acquiring or recovering a competent knowledge of the latin language*. Philadelphia, United States: Joseph Whetham.

Smiley, A. H., and Fisher, M. (2022). The golden age is behind us: how the status quo impacts the evaluation of technology. *Psychol. Sci.* 33, 1605–1614. doi: 10.1177/09567976221102868

Swingewood, A. (1977). *The myth of mass culture*. Macmillan International Higher Education.

The Waltz (1843). *The New World*, 150–152.

Trzesniewski, K. H., and Donnellan, M. B. (2014). "Young people these days…" evidence for negative perceptions of emerging adults. *Emerg. Adulthood* 2, 211–226. doi: 10.1177/2167696814522620

Twenge, J. M., Haidt, J., Joiner, T. E., and Campbell, W. K. (2020). Underestimating digital media harm. *Nat. Hum. Behav.* 4, 346–348. doi: 10.1038/s41562-020-0839-4

Twenge, J. M. (2017). iGen: Why today's super-connected kids are growing up less rebellious, more tolerant, less happy–and completely unprepared for adulthood–and what that means for the rest of us. *Simon and Schuster*.

Wartella, E. A., and Jennings, N. (2000). Children and computers: new technology. Old concerns. *Fut. Child.* 10, 31–43.

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 20 October 2023
DOI 10.3389/fpsyg.2023.1159999

OPEN ACCESS

EDITED BY
Andrea Szymkowiak,
Abertay University, United Kingdom

REVIEWED BY
Danielle Sulikowski,
Charles Sturt University, Australia
Andree Hartanto,
Singapore Management University, Singapore

*CORRESPONDENCE
Leopold Helmut Otto Roth
✉ leopold.roth@univie.ac.at

RECEIVED 06 February 2023
ACCEPTED 02 October 2023
PUBLISHED 20 October 2023

CITATION
Kaminger S, Roth LHO and Laireiter A-R (2023) #Blessed: the moderating effect of dispositional gratitude on the relationship between social comparison and envy on Instagram.
Front. Psychol. 14:1159999.
doi: 10.3389/fpsyg.2023.1159999

#Blessed: the moderating effect of dispositional gratitude on the relationship between social comparison and envy on Instagram

Stella Kaminger[1], Leopold Helmut Otto Roth[2]* and Anton-Rupert Laireiter[1,3]

[1]Faculty of Psychology, Institute for Clinical and Health Psychology, University of Vienna, Vienna, Austria, [2]Faculty of Psychology, Institute of Department of Occupational, Economic, and Social Psychology, Motivation Psychology, University of Vienna, Vienna, Austria, [3]Division of Psychotherapy and Gerontopsychology, Department of Psychology, University of Salzburg, Salzburg, Austria

Introduction: The role of dispositional gratitude as a positive psychological resource and prosocial personality trait in real life interactions militates in favor of its introduction to the research field of social media.

Methods: Based on a literature review of the previously studied relationship of dispositional gratitude with social comparison and envy in offline settings, a twofold moderation model was proposed and quantitatively tested in a cross-sectional sample of N = 268 Instagram users aged between 18 and 40 years. Additionally, the dual conceptualization of benign and malicious envy was scrutinized by validating its respective connections with affective outcomes and inspiration on Instagram.

Results and discussion: Dispositional gratitude serves as a protective factor when using Instagram by significantly mitigating the relationship of social comparison and malicious as well as general envy on Instagram. Furthermore, the results support the more nuanced understanding of envy as a dual construct in the face of social media use.

KEYWORDS

dispositional gratitude, social comparison, benign envy, malicious envy, social media, Instagram

1. Introduction

Blog posts with titles like "#Blessed: Why it's not gratitude if you need to advertise it on social media" (Theresa Christine, 2014), "How to practice gratitude without feeling like an Instagram cliché" (Hoff, 2019), and "How gratitude fixed my relationship with Instagram" (Coyne, 2018) indicate the relevance of positive-psychological research for the area of social media. As noted in previous work, social media use can exhibit both positive and negative characteristics with potentially favorable as well as harmful effects on the user's well-being (Fox and Moreland, 2015; Radovic et al., 2017; Verduyn et al., 2017). Therefore, it is essential to explore the circumstances and mechanisms of both outcomes to promote future beneficial usage behavior. Especially, regarding the increasing number of social media users and the

related psychological costs, variables from Positive Psychology show the potential to function as protective factors when using social media (Pantic, 2014; Hansen, 2015; Winata and Andangsari, 2017). For this purpose, the current paper proposes the implementation of gratitude into the research field of social media by exploring its association with social comparison and envy.

1.1. Gratitude and social media

Whereas state gratitude refers to the emotional response to a blessing received, trait gratitude considers a habitual and constant characteristic and tendency to experience gratitude with higher frequency and intensity (Wood et al., 2008). Significant to a broader conception of dispositional gratitude is the propensity to appreciate consciously the positive aspects of life and to consider one's own abilities and possessions as valuable gifts (McCullough et al., 2002; Wood et al., 2010; Winata and Andangsari, 2017). Regarding the aim of integrating gratitude into the research context of social media, the social component of gratitude is especially relevant. To experience as well as to express gratitude strengthens interpersonal relationships and functions as a social engine and amplifier of prosocial behavior (McCullough et al., 2008; Blickhan, 2018). Inherently, gratitude is considered an intrinsic value and the counterpart of extrinsic demands such as materialism, fame, and possession (Froh et al., 2011). There is a growing body of research on the moderating effect of dispositional gratitude in the context of social variables. In this light, dispositional gratitude was found to downsize negative health consequences associated with a lower socioeconomic status (Hartanto et al., 2018). Dispositional gratitude further serves as a moderator when looking at affective needs and victimization frequency in romantic interpersonal relationships (Griffin et al., 2016). From a meta-analytic point of view, dispositional gratitude is described as an especially meaningful trait in predicting well-being (Portocarrero et al., 2020). Despite the positive effects of gratitude in a variety of domains, its implementation in the research context of social media so far is rather sparse. The existing evidence indicates that the potential of dispositional gratitude as a resource in offline-contexts can be transferred to online-interactions. The reciprocity of gratitude, for instance, applies to an online-context too and is present toward contacts that already existed offline (Godawa et al., 2019). Social media users with higher dispositional gratitude undergo less social comparison and exhibit an overall lower tendency to compare with others socially (Winata and Andangsari, 2017). Moreover, a gratitude intervention on Instagram was found to enhance the users' dispositional gratitude. Due to its interactive features and the social mechanisms behind sharing pictures, Instagram is considered a platform potentially reinforcing gratitude (Koay et al., 2020). On the downside, the platform focuses on aesthetic enhancement by encouraging the use of filters, which in turn favors a strong positivity norm and impels users to mainly share positive and optimized aspects of their lives. As a result, Instagram promotes passive usage habits and almost continuous social comparison processes (Lup et al., 2015). By fostering social comparison, social media use further relates to feelings of envy and decreased well-being (Krasnova et al., 2013; Lange and Crusius, 2015; Verduyn et al., 2015).

1.2. The concept of envy

Generally, envy refers to the painful experience of the perceived superiority of others and is considered a frequent cause of dissatisfaction and decreased well-being, as it results from aligning one's self-worth with the comparison to others (Smith, 2007). Theoretically, two conceptualizations of envy can be differentiated: in the *unitary approach*, envy genuinely entails feelings of pain and hostility and yet possibly triggers positive as well as negative reactions. To do justice to this twofold characteristic of envy, the *dual approach* distinguishes two different forms of envy. Accordingly, *benign envy* aims at reducing the superiority of the envied person by motivating them to catch up, whereas *malicious envy* does so by downgrading, disparaging, and harming the envied (Crusius et al., 2020). Therefore, benign envy is accompanied by active emulation and malicious envy comes along with aggression (Lange et al., 2018). Empirical research as well as factor analyses reveal two inherent factors with weak intercorrelations ($r_s = -0.07$ to 0.32; Lange and Crusius, 2015), hinting at the existence of two forms of envy (Crusius et al., 2020).

To come to terms with this recent debate, we define envy here in a general form as well as in its dual specification for a more differentiated perspective. As previously noted, and generally implemented in social media research, the consideration of the envy subtypes permits an in-depth analysis and reflects the different social media usage outcomes as potentially harmful or beneficial for the users (van de Ven, 2016; Wu and Srite, 2021). Benign envy induces positive outcomes such as inspiration (Meier and Schäfer, 2018), elevated goal setting, better performance (Lange and Crusius, 2015) and positive affect (van de Ven et al., 2009; Braun et al., 2018; Meier and Schäfer, 2018). Malicious envy is associated with "schadenfreude" (Lange et al., 2018), counterproductive work methods (Braun et al., 2018), hostility (Crusius et al., 2020), and negative emotions (Braun et al., 2018). These connections are not applicable to the other form of envy, respectively (Crusius et al., 2020).

When transferring the dual understanding of envy to the research area of gratitude, the findings emphasize the complexity of envy as a psychological phenomenon: Gratitude is defined as a counterpart of and an incompatible characteristic to envy (McCullough et al., 2002). The dual conception of envy allows a more detailed view, with gratitude being a positive predictor of benign and a negative predictor of malicious envy (Xiang et al., 2018).

1.3. The current study

Considering that social media use contributes on the one hand positively and on the other hand negatively to users' well-being, it is especially relevant to expand our knowledge on traits and behaviors inhibiting negative usage habits and promoting positive outcomes (Hansen, 2015). Dispositional gratitude might help to gain new insights into individual usage styles. The relationship between gratitude and social comparison as well as envy suggests that dispositional gratitude might be decisive in whether social media use leads to positive rather than negative consequences. To

test this assumption empirically, a twofold moderation model is proposed.

Current research shows that both types of envy are closely associated with social comparison processes (Meier and Schäfer, 2018). Dispositional gratitude is related negatively to the social comparison orientation of social media users (Winata and Andangsari, 2017) and in addition, serves as a positive predictor of benign and as a negative one of malicious envy (Xiang et al., 2018). Taken together, these findings strongly support the assumptions of our moderator model. We therefore assume that dispositional gratitude positively moderates the relationship of social comparison and benign envy and negatively that of social comparison and malicious envy. As a complementary *post hoc* analysis, we explore the moderating effect of dispositional gratitude on the relationship of social comparison and general envy on Instagram. In the context of our model, social comparison processes can operate in two ways, as a general and a specific phenomenon. The first one targets a common trait and the second one relates to Instagram use in specific. For validating the dual conception of envy, we look at their relationships with inspiration on Instagram and positive and negative emotions, supposing that benign envy positively predicts inspiration and positive affect and malicious envy shows no relationship with inspiration and a positive one with negative affect.

2. Materials and methods

2.1. Sample

The required sample size for detecting a small to medium effect ($f^2 = 0.1$) with a given power of 0.8 was computed using G*Power 3.1 (Faul et al., 2009), which recommended a minimum number of 114 individuals. In view of range restriction in moderation analyses and in orientation toward existing studies, we collected *ad hoc* data of 303 Instagram users, which resulted in a sample size of 268 Instagram users after the exclusion of incomplete datasets and four cases which completed the survey in less than 4 min. The observed power for each model was computed and can be found on OSF.[1] Following the inclusion criteria of young and emerging adulthood (18–40 years; Berk, 2011), our sample shows a mean age of 22.84 years (SD = 3.85, min. = 18, max. = 39). Addressing participants in this younger age segment ensured a representation of Instagram users that are especially concerned with self-branding and impression-management, which in turn makes social comparison processes more prevalent (Alfasi, 2019). Further recruitment criteria included people at least occasionally using Instagram and exhibiting sufficient German language skills for questionnaire handling. The majority of the resulting sample identified as female (63.05%) followed by male (36.19%) and two participants identifying as diverse (non-binary; 0.74%). We did not distinguish between gender and sex and did not measure the alignment between the two in this study. Future studies, aiming to target sex and/or gender differences regarding our hypothesis should include the respective measures to allow respective comparisons. Additional descriptive data on the sample is presented in **Table 1**.

TABLE 1 Descriptive details on sample ($N = 268$).

Country	
Austria	73.13%
Germany	20.14%
Other	6.73%
Relationship	
Single	50.00%
Relationship	50.00%
Education	
No high school	5.22%
High school	69.40%
University	25.38%
Followers (median)	300
Following (median)	250

2.2. Measures

To test our hypotheses, we collected data from each participant with six psychometric scales. Gratitude was measured using the *Gratitude Questionnaire* (GQ-6; McCullough et al., 2002) in its German adaptation (Samson et al., 2011), using a seven-point Likert scale (1 = *strongly disagree*, 7 = *strongly agree*). It showed acceptable reliabilities in earlier studies (McCullough et al., 2002: α = 0.82; Samson et al., 2011: α = 0.73).

To record individuals' general tendency for social comparison, we used the 11-item *Iowa-Netherlands Comparison Orientation Measure* (INCOM; Gibbons and Buunk, 1999) in its German version by Schneider and Schupp (2011). Social comparison tendency was assessed on a five-point Likert Scale (1 = *I disagree strongly*, 5 = *I agree strongly*), with a reliability (Cronbach's α) ranging from 0.78 to 0.85.

For registering social media specific comparison processes we added the six-item *Social Comparison Tendency on Facebook Scale* (COM-F; Steers et al., 2014), translated into German and modified for Instagram by Meier and Schäfer (2018). It is referenced here as *COM-I* (Cronbach's α = 0.85; Steers et al., 2014), using a five-point Likert scale as well.

Additionally, participants completed the *Benign and Malicious Envy Scale* (BeMaS; Lange and Crusius, 2015) for benign (α = 0.84–0.90) and malicious (α = 0.84–0.91) envy subtypes and general envy by calculating the total score. The envy scales were scored using a five-point Likert Scale (1 = *I disagree strongly*, 5 = *I agree strongly*). To gather Instagram specific envy, we applied the respective German adaptation by Meier and Schäfer (2018).

For registering inspiration, we used the *Inspiration Scale* (Thrash and Elliot, 2003) in its German adaptation (Meier and Schäfer, 2018) on a five-point Likert scale, showing Cronbach's α's between 0.90 and 0.95 (Thrash and Elliot, 2003).

As our final psychometric measure, we applied the widely used *Positive and Negative Affect Schedule* (PANAS; Watson et al., 1988;

1 https://osf.io/qm9g5?view_only=d7596b60e9f648438612474233a97f78.

German version by Breyer and Bluemke, 2016) with α = 0.86 for both subscales. Again, a five-point Likert scale was used (1 = *very slightly or not at all*, 5 = *extremely*) to assess positive and negative affect with ten items each.

2.3. Procedure

Mainly, participants were recruited via social media, where an invitation and the link to the online-study were spread. No compensation for participation was offered. First of all, participants were asked to give informed consent for participation. Only then could they continue with filling out the respective questionnaires. The measures were presented in a fixed order to prevent differences in mutual interference of the measures for each participant. Contact details were presented in case any questions arose.

2.4. Data analysis

We excluded all participants with incomplete data or overly fast completion time (4 min). From 303 initially sampled individuals, 268 remained in the final sample. We initially calculated descriptive statistics and Cronbach's α reliability coefficients for all scales and subscales used. To deliver a general oversight on the relationship between measurements, we obtained all correlations (see text footnote 1). To conduct the planned analysis, we tested the predictability of social orientation on envy and its subscales (benign and malicious) as well as the moderating role of gratitude on the relationship. This was done, using the general social orientation measure (INCOM) as well as the social media specified measure (COM-I). This procedure resulted in six models, which are described in **Tables 4**, **5**. Wherever significant, the moderation is plotted, using simple slope analysis (**Figures 1** (3). Last, we tested the consequences of benign and malicious envy on inspiration as well as positive and negative affect. Note that the term prediction is used non-causal in this article as the data does not allow for immediate causal conclusions. Rather, the term is used to describe explained variance. All reported regression coefficients are z-standardized.

2.5. Open science statement

The current study was not pre-registered, yet we follow the open science recommendations of open data and open code, which can be obtained from see text footnote 1.

3. Results

3.1. Descriptive statistics, reliabilities, and social media usage

Table 2 summarizes the descriptive statistics of the applied scales as well as their coefficient alphas. Apart from the GQ-6, all scales showed satisfactory reliability. The alpha of the scale states a limitation to the presented analysis.

TABLE 2 Descriptive details on applied scales.

	M	SD	Median	α
GQ-6	5.92	0.73	6.00	0.66 [0.61, 0.72]
INCOM	3.19	0.48	3.27	0.82 [0.78, 0.85]
COM-I	2.23	0.80	2.33	0.78 [0.74, 0.82]
Envy (general)	1.85	0.64	1.80	0.84 [0.81, 0.87]
Envy (benign)	2.21	0.94	2.20	0.87 [0.85, 0.89]
Envy (malicious)	1.49	0.62	1.20	0.79 [0.75, 0.83]
Inspiration	3.04	1.04	3.25	0.92 [0.91, 0.94]
Positive affect	3.19	0.63	3.30	0.83 [0.79, 0.86]
Negative affect	1.97	0.59	1.90	0.81 [0.77, 0.84]

TABLE 3 Posted content categories by frequency in percentage (%).

	Almost never				Almost always
Fitness	77.61	11.19	7.08	2.61	1.49
Food	59.70	17.91	13.80	7.46	1.11
Fashion	86.94	5.97	4.85	1.49	0.74
People	18.65	14.92	16.41	25.74	24.25
Travel	17.91	8.20	20.52	33.95	19.40
Celebrities	91.41	7.08	1.11	0.37	0.00

Analysis of broad items regarding social media usage showed that 47.01% indicated to use it to gather information, 62.68% to interact with others, 88.06% for entertainment, and 49.63% for inspiration or motivation. Further items included distraction (54.47%), self-presentation (19.40%), keeping up-to-date (41.41%) as well as being creative (22.38%). The wording of items on additional information of Instagram usage style was chosen following Meier and Schäfer (2018).

Participants have been asked how much time they spend on Instagram per day (subjective estimation) which was answered by 265 participants. A total of 10.18% indicated 10 min or less, 24.53% reported 11–30 min, 33.96% said 31–60 min, 21.50% 1–2 h, 9.81% 2–3 h, and 0.02% more than 3 h per day.

As summarized in **Table 3**, participants were most likely to post self-related content (people and travel). The category *people* included selfies, group pictures, and portraits.

3.2. Part I: predictability of envy and its facets

Following the planned analysis, overall envy as well as its subscales was predicted by the general tendency for social comparison. This prediction was significant and positive in all three models, indicating that social comparison tendencies relate toward higher experiences of envy (**Table 4**). For general and malicious envy, gratitude is negatively related to envy experiences, indicating a reversed relationship against social comparison tendencies. For malicious envy, social comparison tendencies and gratitude interacted on their effect on envy experience.

The simple slope analysis in **Figure 1** indicates that individuals with higher mean gratitude (full-line) experience less malicious envy than individuals with less gratitude (fine-dotted line).

TABLE 4 Prediction of envy by general social comparison tendencies (INCOM).

	Envy					
	General		Benign		Malicious	
	β	CI	β	CI	β	CI
Intercept	0.01	−0.11, 0.12	0.01	−0.11, 0.12	0.01	−0.11, 0.13
INCOM	**0.26**^	**0.14, 0.38**	**0.28**^	**0.16, 0.40**	**0.13***	**0.00, 0.25**
GQ-6	**−0.18**$^{+}$	**−0.30, −0.06**	−0.08	−0.20, 0.04	**−0.25**^	**−0.37, −0.13**
Interaction	−0.07	−0.16, 0.01	−0.05	−0.13, 0.04	**−0.09***	**−0.17, −0.00**
R^2 (adjusted)	0.100		0.081		0.072	

*$p < 0.05$, $^{+}p < 0.01$, $^{\wedge}p < 0.001$. Bold values indicates significance.

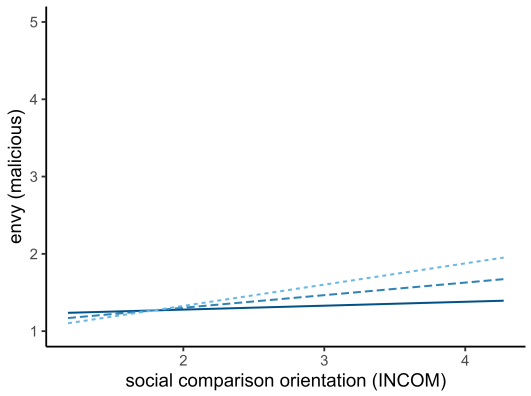


FIGURE 1
Simple slope analysis of gratitude on general social comparison. Full line (−1 SD) and fine-dotted line (+1 SD).

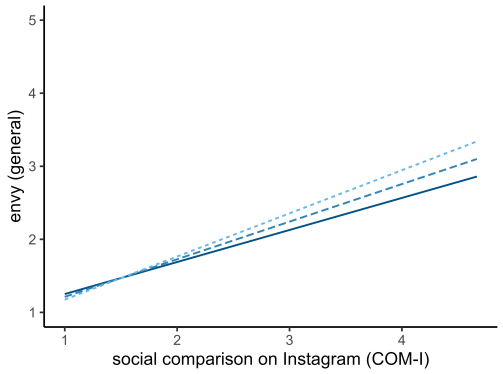


FIGURE 2
Simple slope analysis of gratitude on social media specific social comparison and general envy. Full line (−1SD) and fine-dotted line (+1SD).

When using the social media specific measure of social comparison (COM-I), we observed a very similar effect in directionality but higher explained variance across the models (Table 5). Again, social comparison predicted all facets of envy positively, yet meaningfully stronger. Gratitude moderated the relationship for general as well as malicious envy in the same directionality as before, where individuals with higher trait gratitude experience less envy (Figures 2, 3). It is noteworthy that there was no observed direct effect of gratitude but solely through the moderation of the effect of social comparison tendencies on envy.

3.3. Part II: consequences of envy facets on inspiration and affect

To test the consequences of envy experience, we investigated the predictive value of benign and malicious envy on scales for inspiration as well as positive and negative affect (Table 6). The estimates underline the divergence of the envy dimensions. While inspiration and positive affect are negatively predicted by malicious envy, it is positively related to negative affect, indicating individuals who experience malicious envy tend to also experience less inspiration, less positive affect and more negative affect. On the other side, benign envy positively predicts inspiration and positive affect and was neutral to negative affect.

4. Discussion

4.1. Summary

In the present study, we investigated the relationship between social comparison tendencies and envy as well as the moderating role of dispositional gratitude. We observed that comparison tendencies and envy are related, on the overall level as well as by envy-facets (malicious and benign). Dispositional gratitude moderated this relationship, depending on the perspective on social comparison tendencies and envy conceptualization. Following, we explored the consequences of benign and malicious envy on subjective experiences of inspiration, positive and negative affect. This illustrates that malicious envy specifically holds potential for negative consequences.

4.2. Dispositional gratitude as a resource for social media use

When faced with the destructive aspects of Instagram use such as social comparison and envy, our findings highlight the potential of dispositional gratitude to serve as a buffer variable in this context. According to our results, dispositional gratitude significantly alters the relationship between social comparison and general/malicious

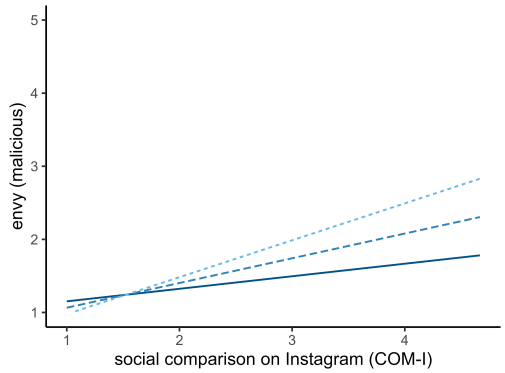


FIGURE 3
Simple slope analysis of gratitude on social media specific social comparison and malicious envy. Full line (−1SD) and fine-dotted line (+1SD).

envy on Instagram. Thus, gratitude can be considered a desirable trait for social media users. Therefore, the importance of promoting dispositional gratitude by interventions is no longer limited to real-life interactions, but transferrable to online-contacts too. Our work paves the way for further studies in this field, as it is crucial to shed light on the circumstances and individual characteristics that determine whether social media use results in favorable or else harmful outcomes. For this purpose, the moderation model proposed can be extended to complementary research questions.

4.2.1. Specificity of social media directed comparison behavior

To our surprise, we observed differing results between the measures for social comparison. The predictive power (see β-values) underlines the stronger effect of social media specific comparison. In addition, only the effect of COM-I on general envy and malicious envy was moderated by individual gratitude. This might illustrate the relevance of context specific measures. As the INCOM measures the general tendency to socially compare oneself in the sense of an underlying traitlike variable, the scale was found to be a subject to socially desirable response behavior (Schneider and Schupp, 2011). The instructions and wording of the COM-I on the other hand refer to more state like reactions to opportunities for social comparison on Instagram specifically. Therefore, our findings form an interesting starting point for further investigations on media-specific measures as well as on the coherence of scales targeting the trait- vs. state-dimension.

4.2.2. Differing effects of envy expression and corresponding effects on other resources

Concerning the continuing debate on how to conceptualize envy, we decided for an interim solution by differentiating between benign and malicious envy subtypes, but also investigating envy as a whole. Still, in the second part of our analysis we aimed at contributing to the discussion by validating the differing envy subtypes via their affective consequences and inspiration on Instagram. Benign envy positively predicted positive affect as well as inspiration on Instagram, while malicious envy positively predicted negative affect and negatively predicted inspiration. In combination with the findings of our moderation models, these results indicate the relevance of the dual envy conception. As our results show, different forms of envy are associated with opposing outcomes in terms of affect and inspiration. Omitting the analysis of envy facets would have resulted in a loss of information. Hence, our study suggests that the dual specification of envy allows an in-depth perspective on the data and enables the derivation of accurate conclusions on the harmful or else beneficial potential of envy. On that basis, not all manifestations of envy are to be considered detrimental ultimately.

TABLE 5 Prediction of envy by social media specific social comparison processes (COM-I).

	Envy					
	General		Benign		Malicious	
	β	CI	β	CI	β	CI
Intercept	−0.01	−0.10, 0.08	0.00	−0.10, 0.10	−0.01	−0.11, 0.08
COM-I	**0.64**^	**0.55, 0.73**	**0.59**^	**0.49, 0.69**	**0.44**^	**0.34, 0.54**
GQ-6	−0.09	−0.18, 0.00	0.01	−0.09, 0.11	**−0.19**^	**−0.29, −0.09**
Interaction	**−0.09***	**−0.18, −0.01**	0.01	−0.08, 0.10	**−0.22**^	**−0.31, −0.13**
R^2 (adjusted)	0.448		0.337		0.316	

*$p < 0.05$, +$p < 0.01$, ^$p < 0.001$. Bold values indicates significance.

TABLE 6 Prediction of inspiration and affect by envy facets.

	Inspiration		Positive affect		Negative affect	
	β	CI	β	CI	β	CI
Intercept	0.00	−0.11, 0.11	−0.00	−0.12, 0.12	−0.00	−0.11, 0.11
Benign	**0.38**^	**0.26, 0.50**	**0.14***	**0.01, 0.26**	−0.01	−0.13, 0.11
Malicious	**−0.16***	**−0.28, −0.04**	**−0.17**+	**−0.30, −0.05**	**0.33**^	**0.21, 0.45**
R^2 (adjusted)	0.122		0.025		0.098	

*$p < 0.05$, +$p < 0.01$, ^$p < 0.001$. Bold values indicates significance.

4.3. Limitations and prospect

To capture the diverse realities of social media users, future research should focus on broader recruiting modalities, extending our approach to a more inclusive sample in terms of gender, age, and education. Generally, self-assessing online-surveys are prone to social desirability responding bias (van de Mortel, 2008). A more experience-based assessment method using online gratitude interventions could potentially lead to less biased responses. Another limitation is the low reliability of the GQ-6 in its German translation. This calls for supplementary validation studies on the GQ-6 in German samples.

Moreover, it has yet to be clarified how envy should effectively be conceptualized and measured. Based on the missing consensus, we see our method of analyzing subscales and the entire scale separately as adequate. Yet, we see that the construction as a two-factor latent construct or other measurement approaches could be valid as well.

To allow for a generalization of our findings, including other social media platforms is of further interest. Future studies in this field should supplement existing research by comparing different social network sites. Furthermore, the statistical model of moderation is inherently correlational and bidirectional, but theoretically deemed causal. Hence, it would be highly desirable for future research to implement an experimental or longitudinal design, as the moderating effect of gratitude is so far exclusively rooted in thorough theoretical derivation of presumed causality (Baron and Kenny, 1986; Wu and Zumbo, 2008).

5. Conclusion

Dispositional gratitude is a key trait variable in the context of social media use and a worthwhile research extension to that field. The high variance explained by the proposed moderation model emphasizes the role of gratitude as a protective factor for Instagram users. Future research should compare the moderating effect of gratitude on social media relevant variables before and after conducting online gratitude interventions. Additionally, the inclusion of clinically relevant variables such as depression or media addiction could be of interest for the extension of our model. By mitigating the relationship of social comparison and envy on Instagram, dispositional gratitude is a valuable resource for social media users. The enhancement of trait-gratitude through social-media-based intervention studies is therefore highly encouraged.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/supplementary material.

Ethics statement

The studies involving humans were approved by the Departmental Review Board, University of Vienna. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

SK conceptualized the study, chose the methodology, collected and analyzed the data, and wrote the original draft as well as the revisions. LR curated and analyzed the data, and wrote and reviewed the manuscript. A-RL supervised the project and reviewed and edited the manuscript. All authors contributed to the article, going through multiple feedback loops and approved the submitted version.

Funding

This research received no grants or funding for the implementation and realization of the study. The open access publication fee was covered by the University of Vienna.

Acknowledgments

The content of the manuscript is based upon SK's thesis at the University of Vienna.

Conflict of interest

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

Publisher's note

References

Alfasi, Y. (2019). The grass is greener on my friends' profiles: The effect of Facebook social comparison on state self-esteem and depression. *Pers. Individ. Differ.* 147, 111–117. doi: 10.1016/j.paid.2019.04.032

Baron, R. M., and Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *J. Pers. Soc. Psychol.* 51, 1173–1182. doi: 10.1037//0022-3514.51.6.1173

Berk, L. E. (2011). *Entwicklungspsychologie*. München: Pearson Studium.

Blickhan, D. (2018). *Positive Psychologie: Ein Handbuch für die Praxis*. Paderborn: Junfermann Verlag.

Braun, S., Aydin, N., Frey, D., and Peus, C. (2018). Leader narcissism predicts malicious envy and supervisor-targeted counterproductive work behavior: Evidence from field and experimental research. *J. Bus. Ethics* 151, 725–741. doi: 10.1007/s10551-016-3224-5

Breyer, B., and Bluemke, M. (2016). *Deutsche version der positive and negative affect schedule PANAS (GESIS Panel). Zusammenstellung sozialwissenschaftlicher items und skalen*. Leibniz: Leibniz-Institut für Sozialwissenschaften. doi: 10.6102/zis242

Coyne, M. (2018). *How gratitude fixed my relationship with Instagram. Verily*. Available online at: https://verilymag.com/2018/09/how-gratitude-fixed-my-relationship-with-instagram (accessed January 17, 2023).

Crusius, J., Gonzalez, M. F., Lange, J., and Cohen-Charash, Y. (2020). Envy: An adversarial review and comparison of two competing views. *Emot. Rev.* 12, 3–21. doi: 10.1177/1754073919873131

Faul, F., Erdfelder, E., Buchner, A., and Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behav. Res. Methods* 41, 1149–1160. doi: 10.3758/BRM.41.4.1149

Fox, J., and Moreland, J. J. (2015). The dark side of social networking sites: An exploration of the relational and psychological stressors associated with Facebook use and affordances. *Comput. Hum. Behav.* 45, 168–176. doi: 10.1016/j.chb.2014.11.083

Froh, J. J., Emmons, R. A., Card, N. A., Bono, G., and Wilson, J. A. (2011). Gratitude and the reduced costs of materialism in adolescents. *J. Happiness Stud.* 12, 289–302. doi: 10.1007/s10902-010-9195-9

Gibbons, F. X., and Buunk, B. P. (1999). Individual differences in social comparison: Development of a scale of social comparison orientation. *J. Pers. Soc. Psychol.* 76, 129–142. doi: 10.1037/0022-3514.76.1.129

Godawa, G., Gurba, E., Rzewucka, P., Tabor, M., and Gurba, K. (2019). "The use of social media and the youth's experience of gratitude," in *Proceedings of the 13th International Technology, Education and Development Conference*, Valencia. doi: 10.21125/inted.2019.0890

Griffin, B. J., Garthe, R. C., Worthington, E. L. Jr., Sullivan, T. N., Larsen, R., Lavelock, C. R., et al. (2016). How positive processes function in negative relationships: Dispositional gratitude moderates the association between affective need and frequency of dating violence victimization. *J. Posit. Psychol.* 11:4.

Hansen, C. (2015). *Gender differences in gratitude, rumination, and social comparison through Facebook (Publication No. 3713944) [dissertation]*. Scottsdale: Northcentral University.

Hartanto, A., Lee, S. T. H., and Yong, J. C. (2018). Dispositional gratitude moderates the association between socioeconomic status and interleukin-6. *Sci. Rep.* 9:802. doi: 10.1038/s41598-018-37109-1

Hoff, V. (2019). *How to practice gratitude without feeling like an Instagram cliché. Yahoo! life*. Available online at: https://www.yahoo.com/lifestyle/practice-gratitude-without-feeling-instagram-060000727.html (accessed January 17, 2023).

Koay, S.-H., Ng, A.-T., Tham, S.-K., and Tan, C.-S. (2020). Gratitude intervention on Instagram: An experimental study. *Psychol. Stud.* 65, 168–173.

Krasnova, H., Wenninger, H., Widjaja, T., and Buxmann, P. (2013). "Envy on Facebook: A hidden threat to users' life satisfaction?," in *Proceedings of the International Conference on Wirtschaftsinformatik, Leipzig, Germany*, Leipzig.

Lange, J., and Crusius, J. (2015). Dispositional envy revisited: Unraveling the motivational dynamics of benign and malicious envy. *Pers. Soc. Psychol. Bull.* 41, 284–294. doi: 10.1177/0146167214564959

Lange, J., Weidman, A. C., and Crusius, J. (2018). The painful duality of envy: Evidence for an integrative theory and a meta-analysis on the relation of envy and schadenfreude. *J. Pers. Soc. Psychol.* 114, 572–598. doi: 10.1037/pspi0000118

Lup, K., Trup, L., and Rosenthal, L. (2015). Instagram #instasad?: Exploring associations among Instagram use, depressive symptoms, negative social comparison, and strangers followed. *Cyberpsychol. Behav. Soc. Netw.* 18, 247–252. doi: 10.1089/cyber.2014.0560

McCullough, M. E., Emmons, R. A., and Tsang, J. (2002). The grateful disposition: A conceptual and empirical topography. *J. Pers. Soc. Psychol.* 82, 112–127. doi: 10.1037//0022-3514.82.1.112

McCullough, M. E., Kimeldorf, M. B., and Cohen, A. D. (2008). An adaptation for altruism: The social causes, social effects, and social evolution of gratitude. *Curr. Dir. Psychol. Sci.* 17, 281–285. doi: 10.1111/j.1467-8721.2008.00590.x

Meier, A., and Schäfer, S. (2018). The positive side of social comparison on social network sites: How envy can drive inspiration on Instagram. *Cyberpsychol. Behav. Soc. Netw.* 21, 411–417. doi: 10.1089/cyber.2017.0708

Pantic, I. (2014). Online social networking and mental health. *Cyberpsychol. Behav. Soc. Netw.* 17, 652–657. doi: 10.1089/cyber.2014.0070

Portocarrero, F. F., Gonzalez, K., and Ekema-Agbaw, M. (2020). A meta-analytic review on the relationship between dispositional gratitude and well-being. *Pers. Individ. Differ.* 164:110101. doi: 10.1016/j.paid.2020.110101

Radovic, A., Gmelin, T., Stein, B. D., and Miller, E. (2017). Depressed adolescents' positive and negative use of social media. *J. Adolesc.* 55, 5–15. doi: 10.1016/j.adolescence.2016.12.002

Samson, A. C., Proyer, R. T., Ceschi, G., Pedrini, P. P., and Ruch, W. (2011). The fear of being laughed at in Switzerland: Regional differences and the role of positive psychology. *Swiss J. Psychol.* 70, 53–62. doi: 10.1024/1421-0185/a000039

Schneider, S., and Schupp, J. (2011). *The social comparison scale: Testing the validity, reliability, and applicability of the Iowa-Netherlands comparison orientation measure (INCOM) on the German population. SOEPpapers on Multidisciplinary Panel Data Research*. Berlin: DIW.

Smith, R. H. (2007). "Envy," in *Encyclopedia of social psychology*, eds R. F. Baumeister and K. D. Vohs (Thousand Oaks, CA: SAGE Publications), 306–307.

Steers, M.-L. N., Wickham, R. E., and Acitelli, L. K. (2014). Seeing everyone else's highlight reels: How Facebook usage is linked to depressive symptoms. *J. Soc. Clin. Psychol.* 33, 701–731. doi: 10.1521/jscp.2014.33.8.701

Theresa Christine (2014). *#Blessed: Why it's not gratitude if you need to advertise it on social media. Elite Daily*. Available online at: https://www.elitedaily.com/life/culture/why-its-not-gratitude-if-you-need-to-advertise-it-on-social-media/595070 (accessed January 17, 2023).

Thrash, T. M., and Elliot, A. J. (2003). Inspiration as a psychological construct. *J. Pers. Soc. Psychol.* 84, 871–889. doi: 10.1037/0022-3514.84.4.871

van de Mortel, T. F. (2008). Faking it: Social desirability response bias in self-report research. *Aust. J. Adv. Nurs.* 25:4.

van de Ven, N. (2016). Envy and its consequences: Why it is useful to distinguish between benign and malicious envy. *Soc. Pers. Psychol. Compass.* 10, 337–349. doi: 10.1111/spc3.12253

van de Ven, N., Zeelenberg, M., and Pieters, R. (2009). Leveling up and down: The experiences of benign and malicious envy. *Emotion.* 9, 419–429. doi: 10.1037/a0015669

Verduyn, P., Lee, D. S., Park, J., Shablack, H., Orvell, A., Brayer, J., et al. (2015). Passive facebook usage undermindes affective well-being: Experimental and longitudinal evidence. *J. Exp. Psy. Gen.* 144, 480–488. doi: 10.1037/xge0000057

Verduyn, P., Ybarra, O., Résibois, M., Jonides, J., and Kross, E. (2017). Do social network sites enhance or undermine subjective well-being? A critical review. *Soc. Issues Policy Rev.* 11:1. doi: 10.1111/sipr.12033

Watson, D., Clark, A., and Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *J. Pers. Soc. Psychol* 54, 1063–1070.

Winata, C., and Andangsari, W. W. (2017). Dispositional gratitude and social comparison orientation among social media users. *Humaniora.* 8:3. doi: 10.21512/humaniora.v8i3.3620

Wood, A. M., Froh, J. J., and Geraghty, A. W. A. (2010). Gratitude and well-being: A review and theoretical integration. *Clin. Psychol. Rev.* 30, 890–905. doi: 10.1016/j.cpr.2010.03.005

Wood, A. M., Maltby, J., Stewart, N., Linley, P. A., and Joseph, S. (2008). A social-cognitive model of trait and state levels of gratitude. *Emotion*. 8, 281–290. doi: 10.1037/1528-3542.8.2.281

Wu, A. D., and Zumbo, B. D. (2008). Understanding and using mediators and moderators. *Soc. Indic. Res.* 87, 367–392. doi: 10.1007/s11205-007-9143-1

Wu, J., and Srite, M. (2021). Envy on social media: The good, the bad and the ugly. *Int. J. Inf. Manage* 56:102255. doi: 10.1016/j.ijinfomgt.2020.102255

Xiang, Y., Chao, X., and Ye, Y. (2018). Effect of gratitude on benign and malicious envy: The mediating role of social support. *Front. Psychiatry* 9: 139.

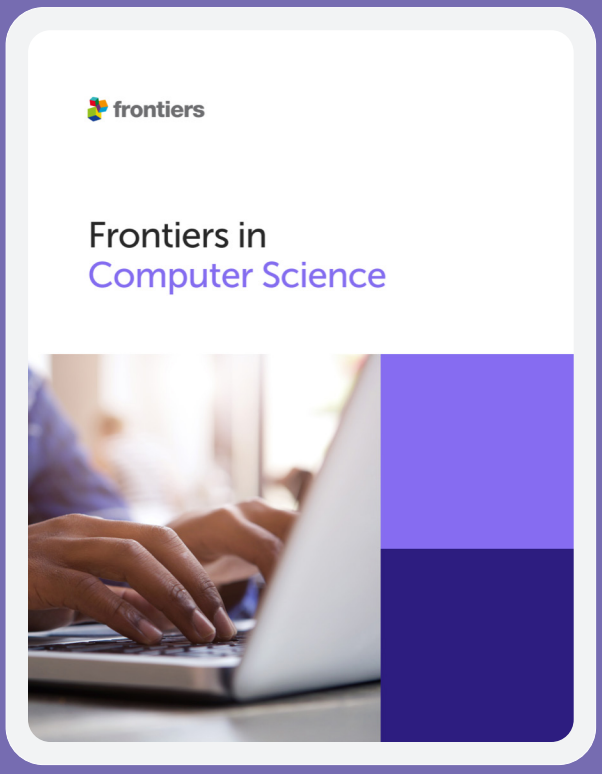
Frontiers in Computer Science

Explores fundamental and applied computer science to advance our understanding of the digital era

An innovative journal that fosters interdisciplinary research within computational sciences and explores the application of computer science in other research domains.

Discover the latest Research Topics

See more →



Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact