Technology in Society 76 (2024) 102444

Contents lists available at ScienceDirect

Technology in Society

journal homepage: www.elsevier.com/locate/techsoc

Why do we not stand up to misinformation? Factors influencing the likelihood of challenging misinformation on social media and the role of demographics

Selin Gurgun [a,*], Deniz Cemiloglu [a], Emily Arden Close [a], Keith Phalp [a], Preslav Nakov [b], Raian Ali [c]

[a] Bournemouth University, UK
[b] Mohamed Bin Zayed University of Artificial Intelligence, United Arab Emirates
[c] College of Science and Engineering, Hamad Bin Khalifa University, Qatar

ARTICLE INFO

Keywords:
Misinformation
Fake news
Social media
Human factors
Social factors
Online silence

ABSTRACT

This study investigates the barriers to challenging others who post misinformation on social media platforms. We conducted a survey amongst U.K. Facebook users (143 (57.2 %) women, 104 (41.6 %) men) to assess the extent to which the barriers to correcting others, as identified in literature across disciplines, apply to correcting misinformation on social media. We also group the barriers into factors and explore demographic differences amongst them. It has been suggested that users are generally hesitant to challenge misinformation. We found that most of our participants (58.8 %) were reluctant to challenge misinformation. We also identified moderating roles of age and gender in the likelihood of challenging misinformation. Older people were more likely to challenge misinformation compared to young adults while, men demonstrated a slightly greater likelihood to challenge compared to women. The 20 barriers influencing the decision to challenge misinformation, were then grouped into four main factors: social concerns, effort/interest considerations, prosocial intents, and content-related factors. We found that, controlling for age and gender, "social concerns" and "effort/interest considerations" have the significant impact on likelihood to challenge. Identified four factors were analysed in terms of demographic differences. Men ranked "effort/interest considerations" higher than women, while women placed higher importance on "content-related factors". Moreover, older individuals were found to be more resilient to "social concerns". The influence of educational background was most prominent in ranking "content-related factors". Our findings provide important insights for the design of future interventions aimed at encouraging the challenging of misinformation on social media platforms, highlighting the need for tailored, demographically sensitive approaches.

1. Introduction

The development of social media has given rise to a new paradigm that influences every facet of society, including instant communication and the sharing of or obtain information across the world. This led to information starting to spread even before its accuracy could be verified, paving the way for the rise of misinformation. Many terms such as, disinformation, fake news and rumours are used interchangeably [1] to refer to incorrect information depending on different factors, such as the format (fake news is typically in a news format), intention (misinformation refers to false information with no intention of harm, whereas disinformation has the intention of deceiving), or the degree of uncertainty about the accuracy of the information (rumours are information with doubt about their accuracy). Throughout this paper, we will use the term *misinformation* as it refers all false and unsubstantiated information that is not supported by evidence or expert opinions [2] regardless of the intention, form or veracity.

Although misinformation is not a recent phenomenon [2,3] concerns over diffusion of misinformation and harmful consequences have increased recently following events such as the COVID-19 pandemic [4] and Ukraine-Russia war [5]. For example, a widespread false information that drinking highly concentrated alcohol may sanitise the body and

* Corresponding author. Poole House, Talbot Campus, Fern Barrow, Poole, BH12 5BB, UK.
E-mail address: sgurgun@bournemouth.ac (S. Gurgun).

https://doi.org/10.1016/j.techsoc.2023.102444
Received 19 July 2023; Received in revised form 21 November 2023; Accepted 6 December 2023
Available online 9 December 2023

kill the COVID 19 virus led to 800 deaths [6] and the inaccurate presentation of the Malaysian Airlines disaster of 2014 as a Ukrainian attack in Russia has contributed to the exacerbation of the Russia-Ukraine conflict [7].

There is a growing body of research that propose solutions to tackle the problem. In this regard, we adopted the categorisation introduced by van der Linden and Roozenbeek [8] to provide an overview of the potential solutions as well as to demonstrate the solution on which our research focuses. The solutions can be divided into four categories: Algorithmic [9–11], legislative [12], psychological [8,13,14] and corrective [15–17]. Our current research primarily focuses on the corrective category, which we have expanded by including corrections made by social media users.

Along with experts [18,19] and fact-checkers [20], social media users also play a paramount role in preventing the spread of misinformation by their corrective attempts. Studies have demonstrated that corrections coming from other users are an effective way to tackle the problem [15,21,22] and as effective as algorithmic corrections [15]. However, data from several studies showed that users on social media are reluctant to correct misinformation they encounter [16,23–27].

The paucity of the behaviour of challenging others who post misinformation poses a significant challenge in combatting misinformation propagation. When individuals refrain from speaking out, it can unintentionally restrict exposure to alternative viewpoints. This can create an environment where misinformation continues to disseminate unchallenged.

Given the significant negative impact of misinformation and the well-established effectiveness of social corrections to combat it, the identification of factors hindering users from challenging misinformation is a valuable endeavour towards mitigating the spread of misinformation. It is also a starting point for enhancing the current design of digital platforms and devising solutions to foster an environment that is more encouraging for social corrections. The implications of this study would inform users' hesitancy factors beyond merely combatting misinformation such as hesitancy and barriers to acting against injustice and prejudice [28–30]. It aims to contribute to Human Computer Interaction (HCI) literature as it one of the first attempts to examine the concept of "online silence" in the context of misinformation and to inform the design of future social media to reduce it.

In this paper, we investigate why people avoid or skip challenging misinformation when they notice in the online environment. We assess the extent to which the barriers identified in Gurgun, Arden-Close, Phalp and Ali [31] apply in the context of social media use. We also grouped these barriers into four factors and investigate the influence of demographics on the likelihood to challenge. We elaborate on the findings and draw implications and provide insights that may inform the design for the future social media platforms where people find it easier and more comfortable to challenge misinformation.

2. Background and motivation

Although the concept of silence has received little attention in the HCI literature, it has been alluded in research on online harassment [32,33], racism [30], cyberbullying [34,35] and online discussions [36,37]. Previous research identified factors that motivate users to respond to misinformation [24,38,39] however, there is limited research exploring the barriers that prevent people from challenging misinformation online.

Given the limited research on the barriers to challenging misinformation online, it might be beneficial to draw on insights from literature regarding silencing in offline environments. Previous studies have explored silence in various contexts and settings, including organisations [40–43], classrooms [44–46] and politics (abstention from politics) [47]. Evidence from the research in offline environments may pave the way for investigating the barriers to challenging misinformation, as it is a well-documented phenomenon in organizational and education literature. While Milliken, Morrison and Hewlin [40] provide insights into the reasons why employees choose to be silent on issues that concern them, Pinder and Harlos [48] developed a conceptual framework regarding employee silence. Other studies have expanded on this framework [41] and tested it in different environments [49] such as aircrews [42].

However, translating the factors from offline to online environments introduces some considerations. For instance, in the online environment, anonymity, invisibility and lack of eye-contact contributes to the negative effects of online disinhibition [50], a phenomenon where individuals feel less restrained and may behave more aggressively due to the lowered perception of social norms [51]. Therefore, it can be expected that people may feel more comfortable due to the lack of visual cues. In a similar vein, research showed that behaviours in certain contexts may not directly correlate to real life behaviours. As an example, research on violent video games showed that in-game behaviour does not necessarily reflect real-life violence [52]. These characteristics of online interactions suggest that barriers for silence or confrontation in offline environments may not be applicable in online context, and thus serve as the motivation for our research.

In the online environment, silence has been defined using various terms and in different contexts. For example, "lurking" is used to refer to the behaviour of observing others and consuming information without actively engaging in online interactions [53,54]. This behaviour however is not inherently related to confrontation; it is a choice to observe without actively participating. The concept of "chilling effects" has been suggested as a form of social conformity which arises when the feeling of being watched limits an individual's choices due to perceived social norms [55]. For the online realm, it refers to restricting one's online self-presentation a result of peer-to-peer surveillance on social media [56]. The "bystander effect" is another term used to describe a behaviour of taking no action such as commenting or reporting or helping a victim in a problematic situation particularly when other people are present [57]. Silence is also explored across various contexts, including online learning environments [58], commenting on online news [59] and within the context of discussing ideas and posting content about political and social issues [60,61] All these terms address the concept of silence in the online environment. However, in our particular context, *silence is defined as the act of refraining from challenging, avoiding responding and being reluctant to counter when encountering false information*. We did not use the terms such as "lurking", "chilling effect" or "bystanders" as they may not comprehensively explain the behaviour of not challenging misinformation but may be influenced by different factors including but not limited to passive browsing, concerns related to self-presentation and a tendency to leave the responsibility to other people. In other words, in our context, silence refers to a deliberate avoidance and reluctance to engage in conversations to discuss, question the veracity of the post or correct it.

The notion of silence in our paper intersects with but also differs from the well-known theory of the "spiral of silence". The spiral of silence theory suggests that people are less likely to express their opinions when they perceive their opinions are in the minority (See Refs. [62,63]). According to this theory, an opinion becoming more prevalent induces a spiralling process in which those who perceive themselves as being in the minority become even more hesitant to speak out, even though, in fact, the minority might be in the majority. The theory posits that the crucial element of the opinion formation is the interplay between personal convictions and the perception of the social environment. This theory is linked to classical theories of social influence where people adjust their behaviour in accordance with the perceived social norms and remain socially accepted [64]. Contrary to the idea that the minority remains silent, minority influence suggests that deviant minorities resist adhering to normative pressures and challenge the majority position [65,66]. Minorities can influence if they promote a coherent and clear social reality that diverges from the majority group. In fact, research indicate that social change is frequently instigated by minorities,

individuals who challenge the status quo [67]. Therefore, our exploration acknowledges that individuals in the minority may exhibit diverse behaviours such as remaining silent or actively advocating their perspectives.

Spiral of silence theory proposes that, this reluctance to speak out is often driven by a fear of isolation when they do not conform to the prevailing opinion. Although studies explored the application of the spiral of silence theory in the context of misinformation and its effects [68,69], our investigation is not specifically focused on settings where users perceive themselves as in the minority. However, some reasons for not challenging misinformation are aligned with this theory such as feeling in the minority ("*I want to avoid being viewed as odd*") and fear of isolation ("*I want to avoid people who posted or endorsed the post isolating me*")

We chose the term "challenging" rather than "correcting" deliberately for this research. Firstly, the term "correcting" presumes that the correction made is the absolute truth, however, which may not always be the case, Sometimes the correction itself can be false information. Secondly, the term "challenging" serves a broader purpose than mere correction. It involves disagreements, questioning or disputes over the post in addition to corrections. Much as the online environment holds the promise of stimulating healthy debates, current social media was not designed as a sphere devoted to rational discussions. The tendency to foster relationships with similar people makes it easier to form "echo chambers" of "filter bubbles" where users are separated from adverse opinions [70]. In fact, in a public sphere like Facebook, people should be able to express their opinions and participate in constructive discussions; however the presence of other members of the "online public" may hinder this interaction [71]. For example, on Facebook, where the entire conversation could be seen by every social contacts, heated interactions and public discussions can be perceived as norm violations [72].

In an environment where the primary goal is not to challenge someone, disagree, or dispute over a content, doing so may be viewed as conflicts. Conflicts emerge when individuals hold opposing thoughts or beliefs on a topic that are perceived as risky and unfavourable in cyberspace [73,74]. There are five ways to handle a conflict including competing, collaborating, compromising, avoiding and accommodating [75]. Although this taxonomy was proposed for offline contexts and primarily within organisational settings, it remains applicable in the online environment. Lee [76] outlines various methods used by social media users to navigate confrontational situations online such as competitive-dominating strategies (e.g., flaming and denouncing), avoidance (e.g., withdrawal), and cooperative-integrating tactics (e.g., apologising, and mediating). Competitive strategies include the act of asserting one's interest to win while the avoiding strategy consists of being silent and avoiding the issue and the person. Among all of them, as a conflict resolution style, cooperative-integrating strategy is the most favourably viewed as it involves compromising, sharing ideas, discussion without evaluation and problem solving. Avoidance on the other hand is the least effective strategy as it does not address the issue [77].

In their work Tagg, Seargeant and Brown [78] discuss a paradox regarding social media users' approach to discussions and their tendency to avoid conflicts. In their study, the participants seemed to hesitate to engage in reasoned debate around different opinions. Instead, they either ignore conflicting opinions or they took measures such as removing content and friends from their feed. In other words, they reconstruct their feed only with the opinions that align with their own. The paradox is that while participants report that Facebook is not suitable for serious discussions, they still use the platform for political expression, but at the same time avoid engaging in controversial topics. This avoidance of expression on Facebook is also in contrast with research revealing that users are less inhibited when expressing themselves to strangers due to the lack of visual cues [51,79]. However on Facebook one could argue that people's offline social roles and responsibilities can influence how they present themselves as they feel pressured to align with societal expectations [80] If users adopt the avoidance strategy as a conflict handling strategy and do not actively engage with opposing views, the platform is unlikely to achieve its goal of facilitating open discussions, allowing users to gain insights from different point of views. This prevents people from gaining a complete understanding of the situation and contributing to a culture of silence, which can in turn make it difficult for people to speak up about issues such as harassment or racism. By identifying the underlying reasons for this silence, we aim to encourage people to adopt cooperative-integrating tactics.

While there are studies examining the user motivations to correct misinformation online [24,26,38], research investigating the barriers that prevent people from challenging misinformation remains limited. A qualitative study by Chadwick, Vaccari and Hall [81] showed that individuals' adherence to the norm of conflict avoidance prevents people from engaging in conversations to correct the misinformation on personal messaging platforms. Another study investigating the obstacles that healthcare professionals face when they correct health misinformation identified three categories intrapersonal (e.g., the lack of time and the perception of limited positive outcomes), interpersonal (e.g., fear of being harassed and bullied), and institutional (e.g., a lack of institutional support and social media training) [82]. These studies focus on health-related topics either within personal messaging platforms such as WhatsApp or involves healthcare professionals using a qualitative approach. While these findings provide valuable insights into the factors influencing the reluctance to challenge misinformation on social media, there may remain several variables, yet to explore in relation to why people may be hesitant. For instance, the level of relationship amongst the users and the level of behavioural privacy on personal messaging platforms differ across platforms [83]. While Facebook is regarded as a semi-public place, WhatsApp is considered to provide a higher level of behavioural privacy [83] and since it is used mainly to connect with people with whom one has strong ties, people feel more comfortable when sharing controversial opinions [84]. Therefore, in this study we aimed to take a broader and exploratory approach by extending the variables beyond those explored in previous research specifically within a semi-public online sphere with acquaintances and with a focus on non-health-related topics using a quantitative method. Demographic variables, such as gender and education, appear to be associated with opinion expression and corrections of misinformation in the online sphere [26]. Based on previous studies, younger individuals were more likely to engage in online participation on political, civic, cultural or health related issues. Those with higher levels of education were also more likely to engage in corrective behaviours related to health issues on social media [26]. However, to the best of our knowledge no research has investigated the relationship between demographic variables and the reasons that prevent users from challenging misinformation.

In sum, while previous studies have focused on the motivation to challenge and have not investigated barriers to challenging misinformation in the online sphere, as well as the roles of factors such as demographics, our study takes a step further and contributes to the field by examining a wider range of social media users beyond healthcare professionals and employing a quantitative approach to identify reasons from different domains regarding the barriers to speaking out.

3. Research questions

To address the aforementioned issues, this research aims to investigate the following research questions.

RQ1 How do social media users engage in challenging misinformation?

RQ1. a What are the various demographic factors (age, gender and education) that might have an influence on the likelihood to challenge misinformation?

RQ2 To what extent are the barriers obtained from literature in different domains applicable to social media users' inaction in challenging others who post misinformation?

RQ2. a How can these barriers be grouped into a smaller set of factors?

RQ3 Which of the identified factor(s) has/have a statistically significant impact on the likelihood to challenge misinformation?

RQ4 What are the differences in the identified barriers based on gender, age and education?

4. Method

4.1. Participants

A total of 250 adults living in the UK were recruited for the study. Participants were compensated for their involvement in completing the survey. Specific inclusion criteria were implemented, which included: 1) being 18 years of age or above; 2) using Facebook with authentic identity; and 3) having encountered misinformation on the platform.

4.2. Data collection procedure

Ethical approval for the research was obtained by the university's research ethics committee. The survey was designed and conducted online using Qualtrics. Participants were provided with an explanation of the research objectives before the survey, and their consent was sought before proceeding. They were also informed regarding the confidentiality of the data, their freedom to participate and the right to withdraw the study as well as their free access to the study's findings. The data collection period was between 31st May and July 7, 2022 through Prolific™ (www.prolific.co), a well-established online platform for recruiting participants for research studies.

4.3. Pilot test

Prior to actual data collection a pilot test was conducted. An online draft questionnaire with the initial set of items was prepared and circulated to students and their network for feedback. 19 participants completed the pilot questionnaire. Several changes were made to make a priori items clearer. In order to assess face validity, we asked participants to elaborate on their thoughts and provide feedback.

4.4. Questionnaire development

The questionnaire included a broad range of questions regarding attitudes and behaviours towards challenging misinformation on Facebook as well as demographic and personal characteristics of the respondents. Facebook was chosen as the platform for this study as it is the most used online social network worldwide with approximately 2.91 billion monthly active users [85] and is used by all age groups [86]. Additionally, 67 % of internet users claimed that they encountered fake news on Facebook [87]. Facebook is also a suitable platform to investigate the interpersonal relations with different social network. On Facebook people are connected with both strong ties (e.g., family, close friends or romantic partners) and weak ties (e.g., former co-workers or neighbours) [88], with the latter dominating Facebook [89]. In his seminal work, Granovetter [90] argued that weak ties are significant because they may serve as bridges to connect network nodes which are more likely to be sources of different and novel information. Weak ties have been linked to exposure to different opinions [91] and people find it more difficult to debate or confront individuals with whom they have a weak connection compared to those with whom they have strong ties [92]. Therefore, the survey questions regarding challenging misinformation focused on challenging or being challenged by an acquaintance on purpose, as it refers to weak ties. It is also a semi-public space where comments and likes are visible to more people than just friends -e.g., if someone comment on a friend's profile that is open to public, everyone will see the comment. Our aim is to explore the concerns users have when challenging misinformation considering factors such as the influence of social dynamics and increased interactions, therefore, instead of a private space we choose a semi-public space. In addition to that, Facebook possesses features like "like", "comment", and "share" which are common and widely used across various social media platforms, the familiarity can enhance participants' understanding and potentially findings may be more broadly applicable.

We provided the definitions of the terms such as "misinformation", "challenging" and "acquaintance" to participants before starting the survey to ensure that all participants have a clear understanding of these terms. Our objective was to mitigate any potential misinterpretation of the survey questions and thereby to enhance the reliability and validity of the data.

4.5. Measures

4.5.1. Demographic characteristics

Participants reported demographic characteristics including age, gender and educational level. Identified gender and age were collected using an open-ended format. Educational level was categorised as primary (compulsory education), further education (vocational training and college), and higher education (university and postgraduate degrees).

4.5.2. Likelihood to challenge misinformation

After providing the definition of challenging misinformation, participants were asked to rate how likely were they to publicly challenge misinformation shared by an acquaintance on Facebook [16,93] on a seven-point scale (1 = Extremely unlikely 7 = Extremely likely) ($M = 3.30$, $SD = 1.94$). We relied on self-reports because although the act of being silent itself is visible, the motivations behind it can be subjected to misinterpretation [94]. Therefore, self-reports were considered a more reliable measure of participants' likelihood to challenge misinformation, as it allowed them to express their thoughts and intentions directly.

4.5.3. Barriers to challenging misinformation

To assess Facebook users' reported barriers for not challenging misinformation we developed a list of 22 items. A practical approach was taken for this objective given the lack of literature and validated measures in this area. Initially, we examined relevant literature, generating items based on previously reported reasons for not challenging misinformation [24,81,82]. In addition to the barriers of these qualitative studies, we kept the scope of the items as broad as possible due to the exploratory nature of this study. Therefore, rather than including items directly related to barriers to challenging misinformation we tried to generate a range of reasons that might be related to online silence. In order to identify unexplored barriers, we searched literature on topics such as organisational behaviour, communication, human-computer interaction (HCI) and psychology to gain insights into why people remain silent in other settings. This resulted in the creation of several additional items. For example, we derived new items based on the framework developed by Milliken, Morrison and Hewlin [40]. The extensive investigation by Gurgun, Arden-Close, Phalp and Ali [31] about the potential barriers behind people's reluctance to challenge misinformation was the basis for our question items.

Given the exploratory nature of this work, our major objective was to gather as many different barriers as possible without making early assumptions. These 22 items which relate to individuals' barriers for not challenging misinformation were assembled into a questionnaire format. An overview of these items is given in Table 1.

Participants were asked the following question: "If you choose not to challenge misinformation shared by an acquaintance, how much do each of the following reasons apply to you?". They were asked to indicate the extent to which the statements applied to them using a seven-point Likert-scale from "strongly not at all applicable" (1) to

Table 1
References for potential barriers of not challenging misinformation.

Item	Reference
1. I want to avoid being viewed as someone causing problems	[40]
2. I want to avoid being viewed as aggressive	[40]
3. I want to protect my image or reputation	[40,95]
4. I want to avoid damaging relationships with people who posted or endorsed the post	[24,40]
5. People won't delete the post they shared anyway	[24,40]
6. Challenging is ineffective in mitigating misinformation	[24]
7. I want to avoid conflict with people who posted or endorsed the post	[24,81]
8. I want to avoid aggressive reactions from others	[82]
9. I want to avoid people who posted or endorsed the post isolating me	[96,97]
10. Others who posted or endorsed the post feel embarrassed	[40]
11. Others who posted or endorsed the post feel offended	[40]
12. Others who posted or endorsed the post will think that they are seen as untrustworthy	[40]
13. It provides no benefit to me	[82]
14. I want to avoid being viewed as odd	[40,96]
15. It is not my responsibility to correct	[16,27]
16. There are not enough tools on social media for corrections	[98]
17. It takes too much time	[99]
18. Other users will correct it anyway	[16,27]
19. It will be a waste of my effort	[40,82]
20. I don't have enough information to correct	[43]
21. The issue addressed in the post is not important	[100, 101].
22. The issue addressed in the post is not relevant to me	[24,99]

Table 2
Demographic characteristics of the participants ($N = 250$).

		N	%
Gender			
	Woman	143	57.2
	Man	104	41.6
	Non-Binary	3	1.2
Age			
	18–24	44	17.6
	25–34	94	37.6
	35–44	67	26.8
	45+	45	18.0
Education			
	Primary education	36	14.4
	Further education	57	22.8
	Higher education	157	62.8

"extremely applicable" (7). Items were randomly shuffled for each participant. In addition, participants were asked an open-ended question which gave them the opportunity to provide any other reasons that were not listed.

4.5.4. Data analysis

The study included both continuous and ordinal data, which was analysed using SPSS software version 28 (IBM SPSS Inc., Chicago, IL, USA). The sample size in this study is 250 which aligns with the sample size estimated using G*Power software [102] (a priori and post-hoc Power analysis are provided on OSF https://osf.io/cr3bd/). Furthermore, drawing from Green's [103]work, a minimum sample size 50 + 8 m, where "m" represents the number of predictors or independent variables in a regression analysis, the minimum sample size considered appropriate for this study would be 98 which is considered sufficient for assessing the impact of the six predictors on the outcome. For correlations, stability is achieved once the sample size reaches 250 [104]. To establish a benchmark, our sample size also aligns with comparable studies investigating the reasons for particular online behaviour and conducting factor analysis [105,106].

Both parametric and non-parametric tests were used as some data was not normally distributed. Descriptive statistics were used to describe the data and report frequencies. The Mann-Whitney *U* test was used to evaluate the difference between two groups, and the Kruskal-Wallis H test was used to observe the differences between more than two groups. Spearman's rank-order correlations were run to measure the strength and direction of the association between two continuous or ordinal variables. A principal component analysis (PCA) was carried out to determine whether the items could be categorised into meaningful factors. The factor scores obtained from PCA were used as dependent variables to examine whether the barriers varied according to the demographic characteristics of the respondents. Ordinal logistic regression was used to determine which factors had a statistically significant effect on the ordinal dependent variable. Independent samples t-tests, and one-way analyses of variance (ANOVA) were used to compare participants' demographic variables and identified factors.

5. Results

5.1. Participant demographics

In total, 250 participants completed the online survey. Data concerning the participants' gender, age and level of education are shown in Table 2. Most respondents were women (57.2 %), and the mean age was 34.8 years (standard deviation (SD) = 10.96). Majority of respondents had at least a university education (N = 157).

5.2. RQ1: To what extent social media users challenge misinformation?

Participants were asked to indicate the platform where they encounter misinformation the most. Facebook emerged as the primary platform where 51 % (166 respondents) reported encountering misinformation. Following Facebook, Twitter and social network sites (as a general term to refer to various platforms) and Instagram were mentioned as the platforms where participants encountered misinformation most frequently. For a detailed breakdown of the counts and percentages of the platforms where participants encountered misinformation, please see Table 3

Most participants (58.8 %) stated that they were unlikely to challenge (Extremely unlikely, unlikely and somewhat unlikely) others on Facebook (see Fig. 1). Participants aged 18–24 had the highest proportion of respondents (75 %) who were unlikely to take the specific actions in the questions. Conversely, participants aged 45 and over had the highest proportion of respondents (57 %) indicating that they were likely to challenge misinformation. Moreover, respondents with a university education or higher were more likely to challenge (Fig. 2). In general, most participants responded that they were unlikely to challenge misinformation.

5.2.1. RQ1.a: Do demographics (age, gender and education) have influence on the likelihood to challenge misinformation?

A Spearman's rank-order correlation analysis revealed a statistically significant positive correlation between likelihood to challenge and age, rs (248) = 0.135, $p < .05$. Older participants were more likely to challenge misinformation than younger participants. Additionally, a Mann-

Table 3
The counts and percentages of the platforms.

Platform	Counts	%
Facebook	166	51
Twitter	48	15
Social Network Sites	37	11
Instagram	25	8
Whatsapp	13	4
Youtube	11	3
Tiktok	10	3
Reddit	6	2
Online forum	5	2
Snapchat, Line, Telegram, Quora	5	2

Table 4
Rotated structure matrix for PCA with varimax rotation of a four-component questionnaire.

Items	Rotated Component Coefficients				
	Component 1	Component 2	Component 3	Component 4	Communalities
SOCIAL CONCERNS (Cronbach's alpha = 0.908)					
8. I want to avoid aggressive reactions from others	**.806**	.136	.067	.114	.686
2. I want to avoid being viewed as aggressive	**.799**	.211	.223	.086	.740
1. I want to avoid being viewed as someone causing problems	**.791**	.222	.223	.195	.763
7. I want to avoid conflict with people who posted or endorsed the post	**.762**	.226	.136	.079	.656
4. I want to avoid damaging relationships with people who posted or endorsed the post	**.738**	.126	.286	.081	.648
3. I want to protect my image or reputation	**.687**	.218	.257	.054	.588
14. I want to avoid being viewed as odd	**.567**	.243	.380	.015	.526
EFFORT/INTEREST CONSIDERATIONS (Cronbach's alpha = 0.815)					
19. It will be a waste of my effort	.214	**.763**	−.050	.107	.642
17. It takes too much time	.099	**.716**	.076	.126	.544
13. It provides no benefit to me	.137	**.705**	.065	.332	.631
6. Challenging is ineffective in mitigating misinformation	.262	**.606**	.055	−.239	.497
5. People won't delete the post they shared anyway	.316	**.602**	−.022	−.176	.493
15. It is not my responsibility to correct	.132	**.599**	.272	.256	.516
18. Other users will correct it anyway	.156	**.499**	.173	.236	.359
PROSOCIAL INTENTS (Cronbach's alpha = 0.897)					
10. Others who posted or endorsed the post feel embarrassed	.255	.067	**.871**	.067	.832
11. Others who posted or endorsed the post feel offended	.365	.053	**.837**	.086	.844
12. Others who posted or endorsed the post will think that they are seen as untrustworthy	.270	.108	**.815**	.006	.749
CONTENT RELATED FACTORS (Cronbach's alpha = 0.756)					
21. The issue addressed in the post is not important	.078	.410	.150	**.786**	.815
22. The issue addressed in the post is not relevant to me	.073	.410	.165	**.774**	.800
20. I don't have enough information to correct	.223	−.173	−.096	**.644**	.504

Note: Major loadings for each item are bolded.

Table 5
The correlation analysis between likelihood to challenge, PCA generated factors and demographics.

	Age	Gender	Social concerns	Effort/interest considerations	Prosocial intents	Content related factors
Likelihood to challenge	.135[a]	0.124	−.246[b]	−.240[b]	−0.094	−0.124
Age		0.070	−.138[a]	0.009	−0.015	−0.039
Gender			−.141[a]	.165[b]	−0.072	−0.097
Social concerns				−0.052	0.041	0.031
Effort/interest considerations				1.000	0.063	0.030
Prosocial intents						−0.029

[a] Correlation is significant at the 0.05 level (2-tailed).
[b] Correlation is significant at the 0.01 level (2-tailed).

Table 6
Results of ordinal logistic regression analysis.

Predictors	B	Wald χ2 (1)	p	Exp (B)	95 % Wald Confidence Interval for Exp (B)	
					Lower	Upper
Gender(Woman)	−0.494	4.310	0.038	0.610	0.383	0.973
Age	0.019	3.098	0.078	1.019	0.998	1.040
Social concerns	−0.538	20.322	<.001	0.584	0.462	0.738
Effort/interest considerations	−0.566	22.597	<.001	0.568	0.450	0.717
Prosocial intents	−0.128	1.242	0.265	0.879	0.702	1.102
Content related factors	−0.184	2.563	0.109	0.832	0.663	1.042
Dependent variable: Likelihood to challenge						

Whitney *U* test showed a marginally significant difference in challenging misinformation based on gender. There were three non-binary participants that were excluded from the analysis. Men reported being more likely to challenge misinformation (Mdn = 3,00) than women (Mdn = 2.00), U = 8499, z = 1.953, $p = .051$. Furthermore, a Kruskal-Wallis test was run to determine if there were differences in challenging misinformation based on educational level. Median likelihood to challenge scores were not statistically significantly different across education groups, $\chi 2\ (3) = 0.141$, p = .932.

5.3. *RQ2: To what extent are the barriers obtained from literature in different domains applicable to social media users' inaction in challenging others who post misinformation?*

Fig. 3 presents the descriptive data on the extent to which the reasons obtained from the literature apply to participants. The answer "I don't have enough information to correct" was the most applicable reason

Table 7
Independent-samples t-tests results for gender-related differences in four factors.

	Women (N = 143) Mean (SD)	Men (N = 104) Mean (SD)	t (df)	p-value	95 % CI of the Difference
Social concerns	0.0755 (1.025)	−0.128 (0.96)	1.582 (245)	0.05	−0.049–0.456
Effort/interest considerations	−0.113 (1.047)	0.166 (0.92)	−2.179 (245)	0.03	−0.532–0.026
Prosocial intents	0.0721 (1.013)	−0.074 (0.963)	1.143 (245)	0.12	−0.105–0.398
Content related factors	0.112 (0.887)	−0.154 (1.118)	2.014 (190.123)	0.04	0.005–0.527

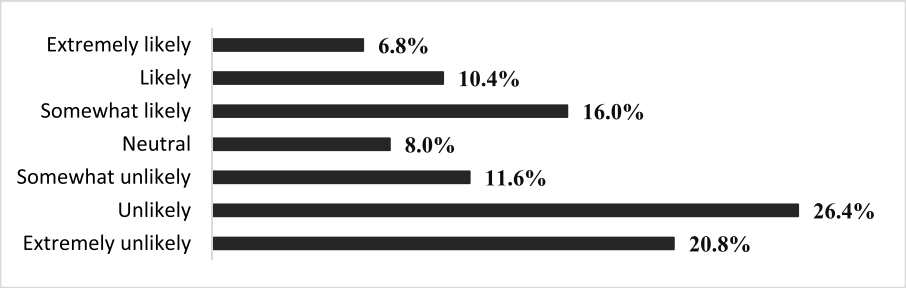

Fig. 1. The likelihood to challenge misinformation.

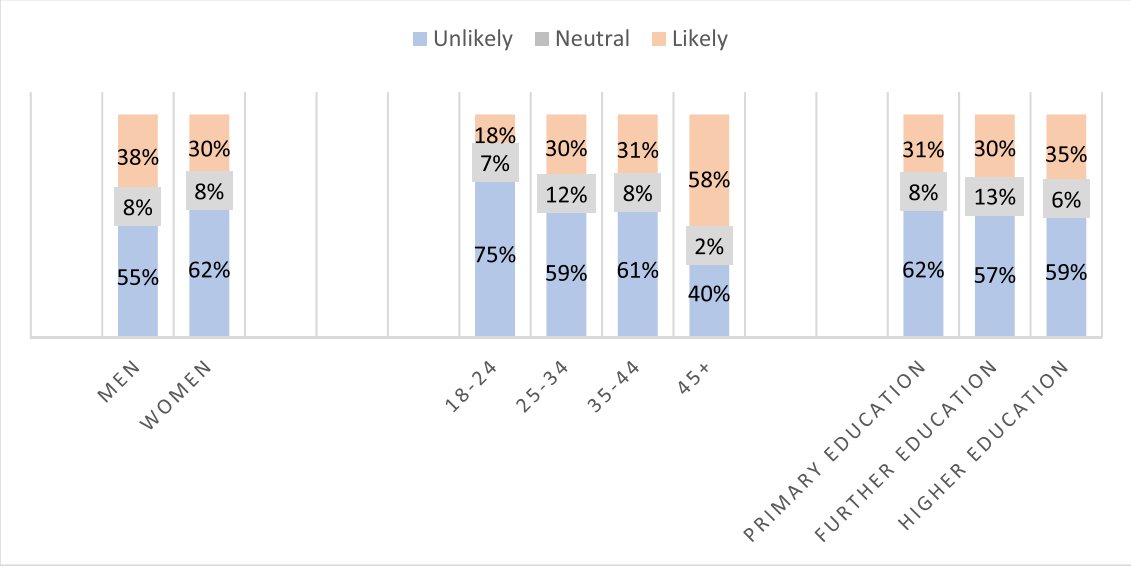

Fig. 2. The likelihood to challenge misinformation based on gender, age and education.

(partly applicable/applicable/extremely applicable) (88 %). The reason "I want to avoid aggressive reactions from others" was the second most applicable barrier reported by the participants (76 %). When asked to list any additional reasons they could think of, as an open-ended question, seven participants mentioned five additional reasons for not challenging misinformation: the person who posts the misinformation is not reasonable or relevant (3), not wanting to come across as a "know-it-all" (2), thinking that the content was posted accidently (1), and seeing that others have already corrected the misinformation (1) (see Fig. 3).

5.3.1. *RQ2.a: Can these barriers be grouped into a smaller set of factors ?*

To determine whether the 22 barriers could be explained by a smaller number of variables, a Principal Component Analysis was performed on 250 total data sets using SPSS®. Table 4 shows the results of a four-factor varimax rotated principal component analysis. We eliminated two items (Item 9 and 16) due to the cross-loading with a difference lower than 0.20 between its primary and alternative factor loadings [107] and low communality (<0.3). The suitability of PCA was assessed prior to analysis. Inspection of the correlation matrix showed that all variables had at least one correlation coefficient greater than 0.3. The overall Kaiser-Meyer-Olkin (KMO) measure was 0.87 with individual KMO measures all greater than 0.7, which is good; or "meritorious" on Kaiser's [108] classification of measure values. Bartlett's test of sphericity was statistically significant (p < .0005), indicating that the data was likely factorisable. Our PCA extracted four components that had eigenvalues greater than one and which accounted for 64 % of the variance. A visual inspection of the scree plot revealed that four components should be retained [109]. Furthermore, a four-component solution met the interpretability criterion. As a result, four components were retained. A Varimax orthogonal rotation was employed to aid interpretability. The factor loadings for the items were 0.40 or above. We label the four factors as follows: Factor 1 (37.1 % of total variance): "social concerns"; factor 2 (12.2 %): "effort/interest considerations", factor 3 (8 %): "prosocial intents" and factor 4 (6.6 %): "content related factors". Component loadings and communalities of the rotated solution are presented in Table 4.

5.4. *RQ3: Which of the identified factor(s) has/have a statistically significant impact on the likelihood to challenge misinformation?*

A Spearman's correlation was conducted to assess the relationship between the four factors, demographics and likelihood to challenge. Preliminary analysis showed the relationship to be monotonic. There was a statistically significant negative correlation between likelihood to challenge and "social concerns" r_s (248) = −0.246, $p < .001$ and between likelihood to challenge and "effort/interest considerations" r_s (248) = −0.240, $p < .001$. This indicates that people who were more likely to challenge were less concerned about these factors. "Social concerns" were negatively correlated with age and gender r_s (248) = −0.138, $p < .05$ and r_s (245) = −0.141, $p < .05$ respectively, indicating as people age social concerns tend to decrease and men are less concerned about their social image or risk of getting in conflict. There was a statistically positive correlation between "effort/interest considerations" and gender r_s (248) = −0.165, $p < .001$, indicating that, men

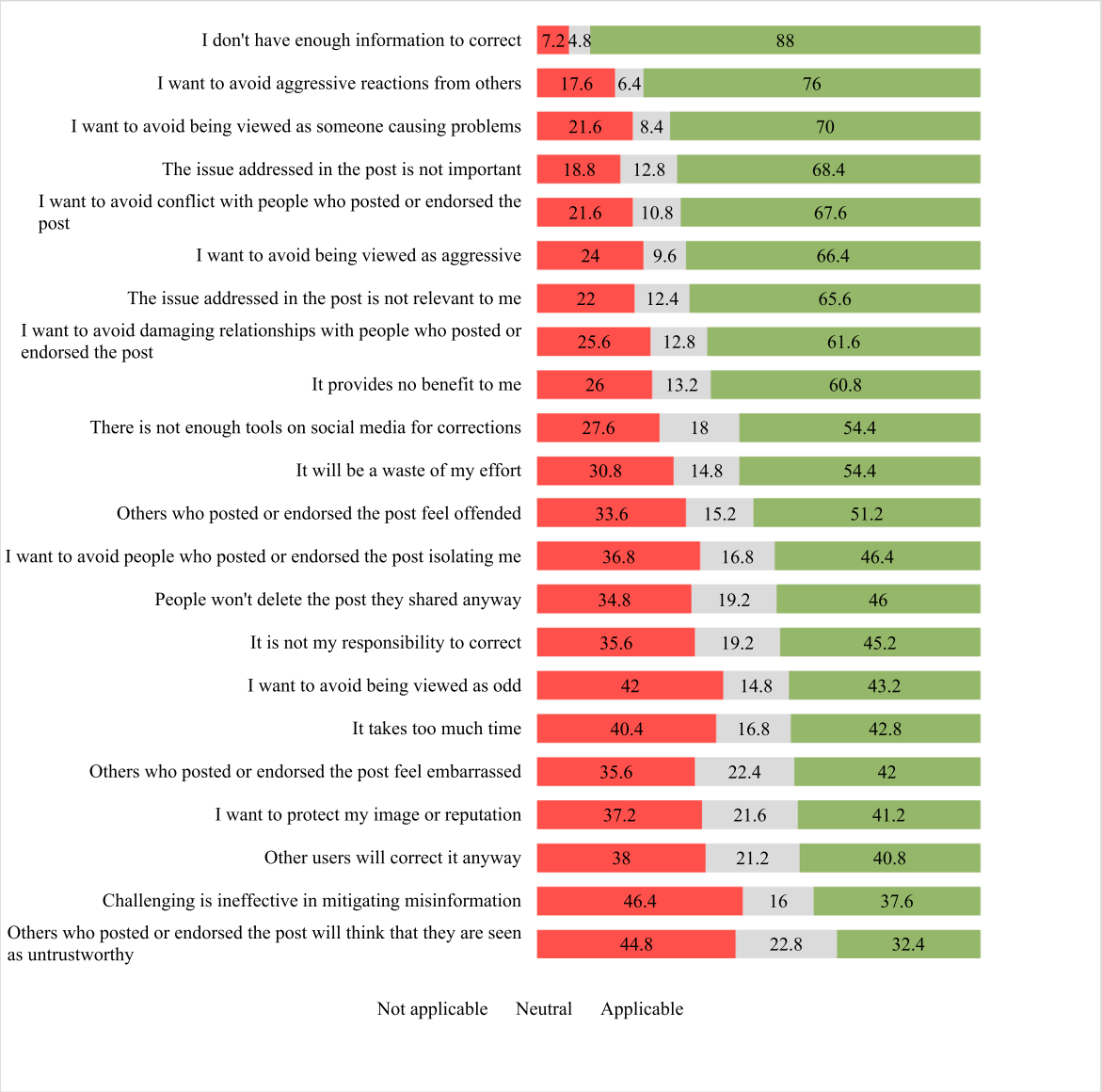

Fig. 3. Frequencies of barriers cited by users for not challenging misinformation on social media (%) (22 items).

were more likely to be concerned about whether their efforts would yield success (see Table 5).

Following the correlation between the variables, a cumulative odds ordinal logistic regression with proportional odds was run to determine the effect of the identified factors (Social concerns, effort/interest considerations, prosocial intents, content related factors) on the likelihood to challenge misinformation controlling for age and gender. In order to investigate the effect, we used the PCA generated factor scores as independent variables. The regression method has been used to generate factor scores. Unlike mean scores which give equal weight to each item and represent the central tendency, this method provides more accurate and meaningful representation of the factor as each item is weighted based on its regression coefficient, taking into account the strength of the relationship between the factor and each item [110].

The assumption of proportional odds was not met, as assessed by a full likelihood ratio test comparing the fit of the proportional odds model to a model with varying location parameters, $\chi 2\ (30) = 77.895$, $p < .05$. Therefore, an examination of the assumption of proportional odds was undertaken by running separate binomial logistic regressions on cumulative dichotomous dependent variables. This examination showed that for most of the variables the assumption of proportional odds appears tenable. The deviance goodness-of-fit test indicated that the model was a good fit to the observed data, $\chi 2\ (1470) = 856.167$, $p = 1$ and the Pearson goodness-of-fit test indicated that the model was a good fit to the observed data $\chi 2\ (1470) = 1534.487$, $p = .118$. The final model statistically significantly predicted the dependent variable over and above the intercept-only model, $\chi 2\ (6) = 52.194$, $p < .001$.

The odds of women considering likelihood to challenge was 0.610, 95 % CI [0.383, 0.973] times that of men, a statistically significant effect, $\chi 2\ (1) = 4.310$, $p = .038$. An increase in "social concerns" was associated with a decrease in the odds of likelihood to challenge, with an odds ratio of 0.584, 95 % CI [0.462, 0.738], $\chi 2\ (1) = 20.322$, $p < .001$. Similarly, a decrease in likelihood to challenge was associated with an increase in the odds of "effort/interest considerations", with an odds ratio of 0.568, 95 % CI [0.450, 0.717], $\chi 2\ (1) = 22.597$, $p < .001$. "Prosocial intents" and "content related factor" did not significantly predict likelihood to challenge as the effect was not statistically significant $\chi 2\ (1) = 1.242$, $p = .265$ and $\chi 2\ (1) = 2.563$, $p = .109$ respectively (See Table 6)

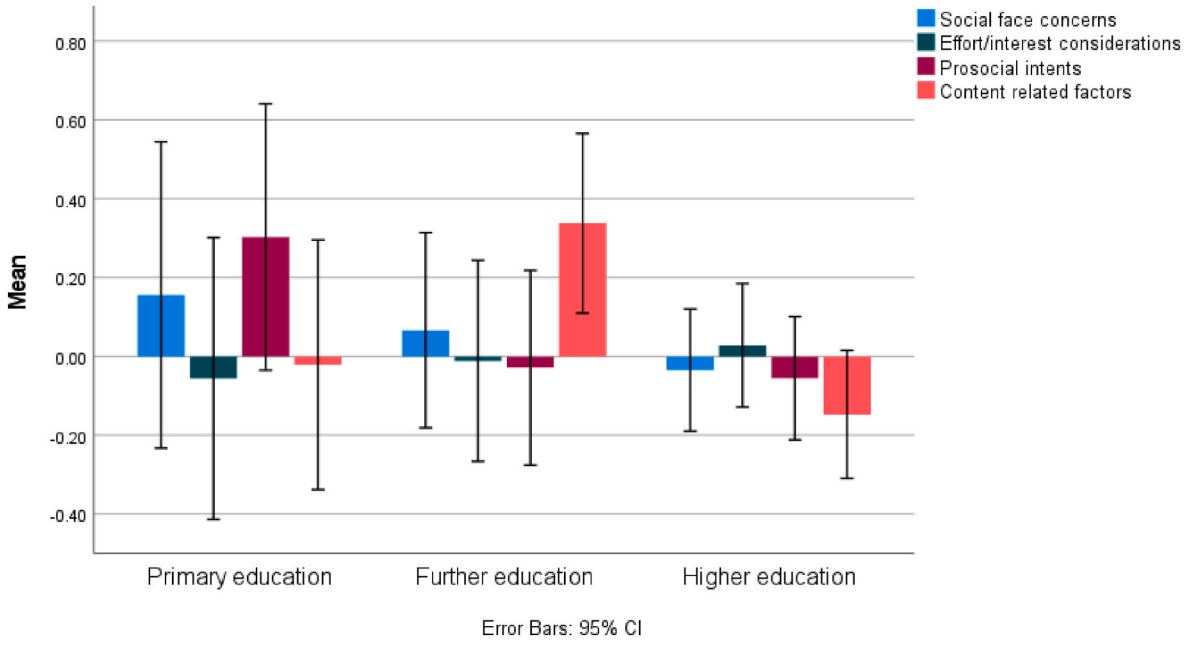

Fig. 4. Analysis of Variance (ANOVA) results for differences in four factors across education levels.

5.5. RQ4: Are there any differences in the identified barriers based on gender and education?

5.5.1. Gender

There were 143 women and 104 men participants. Independent-samples t-tests were run to determine if there were any differences in the four factor scores (Social concerns, effort/interest considerations, prosocial intents, content related factors) between men and women. For all factors there were no outliers in the data, as assessed by inspection of a boxplot. Scores for each gender were normally distributed, as assessed by a Normal Q-Q Plot. The assumption of homogeneity of variances was violated for the "content related factors", as assessed by Levene's test for equality of variances ($p = .019$). However, for "social concerns", "effort/interest considerations" and "prosocial intents" there was homogeneity of variances, as assessed by Levene's test for equality of variances ($p = .563$, .082, 0.968 respectively).

Data are mean ± standard deviation, unless otherwise stated. Scores for the "social concerns" were marginally higher for women (0.075 ± 1.025) than men (−0.128 ± 0.96), but the difference of 0.203 (95 % CI, −0.049 to 0.456) was not statistically significant t (245) = 1.582, $p = .11$. For "effort/interest considerations", scores were lower for women (−0.113 ± 1.047) than men (0.166 ± 0.92), a statistically significant difference of −0.279 (95 % CI, −0.532 to −0.026), t (245) = −2.179, $p = .03$ $d = -0.28$. For "prosocial intents", scores were slightly higher for women (0.072 ± 1.013) than men (−0.074 ± 0.963), but the difference was not statistically significant (mean difference = 0.146, 95 % CI, −0.105 to 0.398), t (245) = 1.143, $p = .12$. Finally, for "content related factors", a Welch *t*-test was run as the assumption of homogeneity of variances was violated. "Content related factors" scores were higher for women (0.112 ± 0.887) than men (−0.154 ± 1.118), a statistically significant difference of 0.266 (95 % CI, 0.005 to 0.527), t (190.123) = 2.014, $p = .04$ $d = 0.26$ (see Table 7).

5.5.2. Education

An analysis of variance (ANOVA) was conducted to evaluate the differences in factor scores across three different education groups. There was one outlier in the data as assessed by inspection of a boxplot for values greater than 1.5 box-lengths from the edge of the box, which was removed from the analysis. Participants were classified into three groups: Primary education (n = 25), Further education (n = 67), and Higher education (n = 157. The data was normally distributed for each group, as evaluated by a Normal Q-Q Plot. Additionally, the homogeneity of variances was maintained, as determined by Levene's test ($p > .05$ for each of the four factor scores).

The results showed no significant differences in means between education groups for "social concerns" (F = 0.542, $p = .58$), "effort/interest considerations" (F = 0.096, $p = .9$), and "prosocial intents" (F = 1.540, $p = .23$). However, for "content related factors", a significant difference in means was found (F = 5.741, $p < .005$). Tukey HSD post hoc tests revealed a significant mean difference in "content related factors" between further education and higher education. Individuals with further education level scored significantly higher on "content related factors" compared to those with higher education, with a mean difference of 0.485 ($p < .05$) (see Fig. 4).

6. Discussion

This study had three aims: (a) to examine the prevalence of 22 barriers to challenging misinformation on social media among Facebook users, (b) to investigate whether these barriers could be classified based on specific factors, and (c) to identify demographic differences in the factors for not challenging misinformation.

While social media appears to be a sphere where individuals feel free to express themselves, literature shows that when users encounter misinformation, they hold back from correcting others [16,23,24,27,81]. Our study aligns with these findings, with the majority of our participants stating a reluctance to challenge misinformation. Facebook was the most mentioned platform among social media users when it comes to encountering misinformation, with over half of our participants acknowledging Facebook as a significant source of misinformation.

6.1. The role of age and gender in challenging misinformation

Our findings showed that older people are more likely to challenge misinformation, which aligns with prior studies suggesting that younger adults typically have fewer discussions and conflicts about politics compared to older adults [74,111,112]. Thorson [113] demonstrates that young people perceive Facebook as too public to engage in political discussions and therefore adopt self-censorship strategies. However, our

findings contrast with previous research about U.S. adults' engagement in COVID-19 misinformation correction which suggested that older adults are less likely to correct others [26]. Several potential explanations might account for these differences. One explanation could be the topics discussed. When the subject is health related, a concern that influences many people, might interact more than older people as they spend more time online [114]. A study on emerging adults showed that more time spent on social media, especially if they are engaged, is associated with increased interactions such as messaging or reacting to contents [115]. Therefore, when younger adults find a topic that engages them, they are more likely to actively interact with the content e.g., by questioning the veracity. Another explanation can be from a cognitive perspective. Although older people spend less time on Facebook, they tend to have more direct interactions when they are on the platform [116]. This direct interaction might be because as individuals get older, they generally become less self-conscious and have fewer experiences of shame, guilt and embarrassment [117], coupled with the fact that self-esteem tends to increase up to the age of 70 [118] which could make them feel more comfortable expressing their thoughts and feelings publicly. Younger individuals, on the other hand, are more active on social media and therefore are exposed to a high volume of information. Studies established that information overload makes it challenging for people to process and critically analyse each piece of information and can lead to an increased avoidance [119,120].

Our results also showed a marginally significant difference in terms of gender. Men showed a slightly greater likelihood to challenge misinformation compared to women. This finding aligns with previous studies suggesting that men are more likely to leave online comments whereas women are more likely to read but not participate in online discussions [121,122]. For example, when women engage with political discussions, they generally do so in a more subtle or less confrontational way relative to men [123]. Even when women encounter online harassment, they choose to ignore it rather than confronting the poster [32]. Although the gender difference is marginal in our study, research suggests that women feel more uncomfortable with confrontations than men. Further research is needed to explore gender differences in online experiences, needs and the barriers to challenging misinformation.

6.2. Barriers to challenging misinformation

People on social media do not often share misinformation with intention to mislead [124]. Indeed, the majority of people said they would need to be paid to share fake news since doing so would put their reputation at risk [125]. If content an individual shares is considered to be fake by others in their network, the sharer could potentially feel embarrassed [126]. Challenging others in this context, would serve as a measure to protect one's reputation. Despite this, people on social media are not often enthusiastic about challenging others. Our study revealed that the major reason participants choose not to challenge misinformation was lack of information regarding the content. This finding is consistent with prior studies regarding the positive relationship between self-efficacy and being confident in one's ability to identify and classify misinformation [127]. Self-efficacy theory highlights the relationship between an individual's belief about their capability to perform a particular behaviour, outcome expectancy and behaviour [128]. Based on this theory, if individuals were not confident in their ability to provide information to challenge misinformation, they would be unlikely to perform the behaviour. Our findings support this proposition. Challenging behaviour is not solely determined by one's intentions or willingness; it is also influenced by the belief that one is capable and equipped enough to do so [128]. Users' perception regarding their belief in their capacity to challenge is highly influential on their decision to challenge misinformation.

Our study identified four factors to challenging misinformation: social concerns, effort/interest considerations, prosocial intents, content related factors. These findings build on a previous study conducted by Gurgun, Arden-Close, Phalp and Ali [31] which identified six factors based on the literature. According to the taxonomy identified by Ref. [31], there were six main reasons: Self-oriented, relationship-oriented, others-oriented, content-oriented, individual characteristics, and technical factors. While we grouped our list into different categories, they were not distinct from these. Reasons identified as self-oriented and relationship-oriented, which included avoidance of conflict or maintaining a positive self-presentation, were grouped under a more general title "social concerns" in our study. Concepts related to efficacy or accountability that were named as self-oriented in Ref. [31] were grouped under the "effort/interest considerations" in this grouping. We did not investigate individual characteristics in this study. Although the categorisation may seem in line with previous study, a different grouping emerged, and technical factors were not heavily emphasised by the users.

Social concerns refer to the concerns people have regarding their behaviour and how it affects their self-image or their relationship with other people. Based on social identity theory, interactions with alternative attitudes (out-group) are often accompanied by greater negative sentiment [129]. Conflict, negative evaluations, or being seen as a troublemaker can threaten one's social identity and lead to potential rejection or exclusion from the group. Given that one of the most prevalent reasons people use social media is social interaction [130], belonging to groups, formation and maintenance of these social connections are crucial for users. Concerns regarding the continuation of social capital leads to social pressure to preserve harmony. Together with the norms of civility, this pressure can influence people to conform to some norms in social media [131]. These norms include, writing positive comments on someone's social media post [132] or not clicking "like" in a response to death-related content [133]. Regardless of how appropriate and required it seems, correcting, challenging, and disagreeing with someone could be regarded as a personal attack [134,135]. Evidence suggests there is a tendency to withhold negative feedback [136,137] and users self-censor their online postings [138]. As people navigate their social networks with the intention of maintaining positive relationships [139,140], they prefer to avoid confronting when they are concerned about negative repercussions of challenging misinformation.

How a person presents oneself to others is an important aspect of social interaction. In order to avoid presenting an unfavourable image, users carefully manage their online identities while being constrained by the expectations of their viewers [141]. In general, people are reluctant to convey negative information due to the discomfort associated with being the messenger of bad news [142]. People strive to present themselves favourably in order to get the approval of others, make a good impression and cultivate a positive self-image [143,144] and the motivation to do so is accompanied by the fear of negative evaluations [145]. To avoid being negatively evaluated (e.g., aggressive, causing problems) people tend to self-silence and not challenge others.

Effort/interest considerations includes the assessment of the effort required, the level of interest individuals have and the expected outcome. In the decision-making process people do a cost-benefit analysis which measures the positive or negative consequences of a behaviour. Protection motivation theory (PMT) [146] can provide insights for why people decide to adopt protective behaviours such as withholding opinions or not engaging with posts. According to this theory, two factors affect individuals' protective behaviour: threat appraisal and coping appraisal. While threat appraisal refers to the severity of the consequences, coping appraisal represents how one respond to the situation. PMT posits that individuals' behaviours are based on a cost-benefit analysis where they weigh the costs of adopting preventive actions against the expected benefits of taking that action. In this case, the precautionary action, remaining silent, may outweigh the expected benefit. According to this theory, if people lack motivation or perceive the effort of challenging as wasteful, challenging misinformation might be seen as requiring too much time and effort. Additionally, if they do

not see direct personal advantages in challenging misinformation, they might be less inclined to invest cognitive effort.

Prosocial intents are related to the intentions that refrain people from challenging misinformation due to concern for potentially harming others in any way. The empathy-altruism hypothesis can provide insights regarding this behaviour. It proposes that individuals who experience empathic concern for others are more likely to engage in selfless acts to help and support them, even when there is no personal gain involved [147,148] In this case, people refrain from informing others that the content they share is incorrect due to desire to avoid embarrassing or offending them, as well as to avoid undermining their trustworthiness. Their prosocial intent arises from a desire to protect others from harm, even if it comes at a cost.

Content related factors such its relevance or importance are also important when deciding to challenge or not to challenge. People do not challenge misinformation when they perceive the content is not relevant or important. This aligns with Petty and Cacioppo [149]'s work on issue involvement which demonstrates that when individuals have a higher level of involvement in a particular issue, the content becomes more influential in shaping their beliefs and important in shaping their cognitive responses. This suggests that, when individuals are involved in a topic they find personally relevant or important, they are more likely to challenge others due to their heightened concern for the accuracy of the information. This factor indicates the importance and relevance of the content as well as users' assessment of whether they have knowledge related to that content to challenge.

Two factors "social concerns" and "effort/interest considerations" had a significant impact on the likelihood to challenge after controlling for age and gender. "Prosocial intents" and "content related factors" did not significantly impact the likelihood of challenging misinformation. Although the item within "content related factors" was mentioned most by the participants as a barrier for challenging misinformation, when combined with other items in the factor, it did not significantly influence likelihood to challenge. One limitation of PCA is that, the factors extracted represent a latent variable that is a combination of the items, meaning the interpretation is based on the underlying construct captured by the factor rather than the individual items [150]. Even though the overall factor did not demonstrate a statistically significant impact, it is important to consider the item that was most mentioned by participants which had a strong influence on their decision to challenge.

6.3. Demographic differences in the factors influencing likelihood to challenge

We investigated potential differences in demographic factors that influence users' likelihood to challenge misinformation on social media. The findings suggested that there is a weak negative correlation between age and "social concerns" such that as individuals grow older, they may become more willing to overcome social concerns and engage in challenging misinformation. As older individuals have more life experiences, they develop a stronger sense of confidence in their beliefs [118] that might make them less concerned about potential social repercussions.

Gender differences were statistically significant in two out of the four factors examined. Women demonstrated lower "effort/interest considerations" and higher scores in "content-related factors" compared to men. Men may place greater importance on assessing the potential outcome or success of their efforts when considering whether to challenge misinformation. They are more concerned about the perceived effort required and whether it will yield the desired result. Studies suggest that men tend to make technology-related choices based on practicality, with their decisions being influenced by the perceived usefulness, especially in the workplace, compared to women [151,152]. This is generally associated with the stereotypical view that men are more "pragmatic and task oriented than women [153]. On the other hand, women appear to prioritise the content itself when deciding to challenge misinformation. They place greater emphasis on factors such as the relevance or importance of the information and having enough information to challenge. One possible reason for this could be that as women are criticised more in online environments [32,121], they try to avoid criticism by deciding whether the subject is relevant to them or if they have knowledge about it before engaging in discussions. These differences underscore the importance of considering gender-specific factors and tailoring interventions to promote effective engagement in challenging false information.

Social norms and perceptions of effectiveness: Societal expectations and cultural norms may also influence men's decision-making process when it comes to challenging misinformation. Men may perceive their efforts as more effective or influential if they believe they have a higher likelihood of achieving the desired outcome. This perception may lead them to place greater importance on assessing the potential outcome or success before challenging misinformation.

Educational background may have a limited impact on users' "social concerns", "effort/interest considerations", and "prosocial intents", but it has a statistically significant influence on evaluation of "content-related factors". People with further education, which could include vocational training or college, show a higher tendency to evaluate content when faced with misinformation compared to people with a university degree and people with primary education. This may be because people with a university degree may feel confident when challenging misinformation as people with higher education tend to have better critical thinking ability [154,155]. Consequently, they do not see content-related factors as significant barriers to challenging misinformation. Conversely, people with primary education may have limited critical thinking skills, which could explain their lack of concern regarding this matter.

6.4. Theoretical and practical implications

This study showed that multiple factors can contribute to individuals refraining from challenging misinformation online. Our findings suggest that addressing the issue of silencing on internet when encountering misinformation requires a comprehensive approach that involves various barriers simultaneously, rather than focusing on a single key element. This adds to the existing literature about self-censorship on social media which focuses on the users' perception of their audience [138,156] and social concerns and technical constraints [157]. Therefore, identifying barriers contributes to a more holistic understanding by considering all relevant factors to make decisions that balance the potential benefits and drawbacks of challenging misinformation, benefiting both users and researchers.

One significant implication of this study is the recognition of social barriers to challenging misinformation. Addressing the widespread issue of misinformation extends beyond mere technical factors such as AI algorithms that creates echo chambers [158,159] and lack of fact checking mechanisms. Social factors also play significant role as barriers in effectively combating misinformation. As our study suggests, there is a prevalent concern among users who feel unsafe when expressing their opinions and stating the facts in the online sphere. Another common belief among users is lack of confidence in the potential success of their efforts. These findings highlight the necessity of an environment where people feel more secure in voicing the facts on social media and a culture that where challenging misinformation is widely accepted.

From a technical perspective, social media platforms can introduce user-centric design strategies, tools and features that can enhance sense of security. Facilitating users' easy access to reliable resources, incorporating easy fact-checking mechanisms and prioritising contents supported by credible sources (e.g. relevant links) can enhance users' confidence in the knowledge they have, empowering them to challenge misinformation more confidently. Rather than solely focusing on increasing engagement, social media platforms' algorithms can prioritise delivering quality content to users. Moreover, users' critical thinking abilities can be improved by incorporating nudges, friction

[160] or prompts. Future research could explore specific design features that enhance users' confidence when challenging misinformation. In order to cultivate a culture where challenging misinformation is not only accepted but actively encouraged, platforms can promote the desired behaviour by recognising and rewarding users who contributes. This can shape the behavioural norms and prevent the fear of repercussions.

Additionally, rather than relying solely on platform-driven solutions, education and media literacy also have the potential to empower users in challenging misinformation [161]. Equipping users with the skills to discern and confront can also empower them to actively participate in conversations.

Prior research efforts aimed at cultivating an environment conducive to constructive dialog. One such example involves integrating a dialog-based online argumentation interface into a comment section of a web page to promote structured argumentation [162,163] This approach aims to improve the quality of online discussions. Scholars also proposed guidelines and principles for fostering effective online discussions in various contexts, such as educational or collaborative platforms [164]. These efforts show the effectiveness of creating interfaces in encouraging users to participate in constructive, fact-based dialogues and improving their information literacy skills. Our study, in that sense, provides insights into the underlying factors behind people's hesitation or reluctance to engage in such discussions that could contribute to design of social media.

Establishing new norms on social media poses a significant challenge. However, making small changes in design, algorithms and education can create an environment where users feel more informed and encouraged to challenge misinformation. This study aims to contribute to the development of future social media designs that can overcome the barriers to challenging misinformation. Such efforts could, consequently, reduce the prevalence of misinformation.

6.5. Limitations and further research

This study has a number of limitations. There are at least three potential threats to the ecological validity of our study. First of all, given the limitations of our sample, it is important to approach the analysis of our findings recognising that they are not nationally representative. Even though our sample size is consistent with previous research recommendations, a larger sample could provide a more thorough understanding and strengthen the reliability of our results. Therefore, it is important to regard this study as exploratory, rather than as a representative sample of the entire population. Secondly, we provided a scenario to participants and asked them to imagine their responses. While this approach allows us to explore reactions, it may lack the authenticity of actual experience. Regarding sample size and representation, our research participants were active Facebook users who encountered misinformation. Therefore, our findings may not be representative of more passive users or users who never encountered misinformation which is a limitation for this study. Lastly, although we specifically relied on self-reports to investigate the reasons as they may not be easily deduced by observation, relying on self-reports may lack the naturalistic and dynamic elements of real-world situations. Participants' responses in a controlled setting might differ from their actual behaviour and motivations in their natural environment. This also could be subject to recall bias [165] or social desirability bias [166] In addition, as this findings were obtained people from the U.K., it may not accurately represent the attitudes and behaviours of other populations and cultures. Previous research suggests that conflict resolution approaches vary between Western and Eastern societies [77]. Therefore, it is important to consider the cultural aspect of this topic. To address this, future research could use observational and experimental designs to assess participants' actual behaviour in challenging misinformation. The barriers examined in this study represent only a limited number of barriers derived from relevant literature. Future research could explore additional barriers and factors. It is also important to recognize that this study focused specifically on Facebook, which may limit the generalisability of the findings to other social media platforms. Different platforms can have different barriers. Future research could explore the differences of the barriers across different platforms. While factor analysis is a useful technique for reducing the number of variables and identifying underlying dimensions, the factors cannot accurately reflect all variables within the factor [150]. Integrating qualitative research methods alongside quantitative factor analysis could provide a deeper understanding of underlying factors.

7. Conclusion

Challenging misinformation on social media is important in order to curb the spread of misinformation, yet there are various barriers that hinder people from doing so. This study highlights the importance of strengthening knowledge and fostering a more harmonious social media environment so that users do not feel restricted when they want to provide facts and challenge incorrect information. This study emphasises the need of recognising and overcoming different barriers by tailoring social media platforms in a way that empowers users and enhances their self-efficacy in challenging misinformation. Addressing these barriers can contribute to the cultivation of social media environment where people feel safe and encouraged to challenge misinformation. The complex interplay of technical and social factors underscores the need for a comprehensive approach to challenge misinformation, one that addresses not just the systems facilitating its spread but also the societal and psychological forces that perpetuate it.

CRediT authorship contribution statement

Selin Gurgun: Conceptualization, Formal analysis, Investigation, Methodology, Resources, Writing – original draft. **Deniz Cemiloglu:** Conceptualization, Methodology. **Emily Arden-Close:** Supervision, Writing – review & editing. **Keith Phalp:** Conceptualization, Supervision. **Preslav Nakov:** Validation, Writing – review & editing. **Raian Ali:** Conceptualization, Methodology, Resources, Supervision, Validation, Writing – review & editing.

Data availability

I have shared the data on OSF. Link was provided on the article.

Acknowledgements

This publication is supported by NPRP-C 14 Cluster Grant # NPRP 14C-0916-210015 from the Qatar National Research Fund (a member of Qatar Foundation). The findings herein reflect the work and are solely the responsibility of the authors.

References

[1] L. Wu, F. Morstatter, K.M. Carley, H. Liu, 'Misinformation in social media: definition, manipulation, and detection', SIGKDD Explor. Newsl. 21 (2) (2019) 80–90.
[2] B. Nyhan, J. Reifler, 'When corrections fail: the persistence of political misperceptions', Polit. Behav. 32 (2) (2010) 303–330.
[3] S. Lewandowsky, U.K.H. Ecker, C.M. Seifert, N. Schwarz, J. Cook, 'Misinformation and its correction: continued influence and successful debiasing', Psychol. Sci. Publ. Interest 13 (3) (2012) 106–131.
[4] M. Cinelli, W. Quattrociocchi, A. Galeazzi, C.M. Valensise, E. Brugnoli, A. L. Schmidt, P. Zola, F. Zollo, A. Scala, 'The COVID-19 social media infodemic', Sci. Rep. 10 (1) (2020) 1–10.
[5] C.Y. Park, J. Mendelsohn, A. Field, Y. Tsvetkov, 'VSaSlov: A Data Set of Russian Social Media Activity during the 2022 Ukraine-Russia War', 2022 arXiv preprint arXiv:2205.12382.
[6] M.S. Islam, T. Sarkar, S.H. Khan, A.-H.M. Kamal, S.M.M. Hasan, A. Kabir, D. Yeasmin, M.A. Islam, K.I.A. Chowdhury, K.S. Anwar, 'COVID-19–related infodemic and its impact on public health: a global social media analysis', Am. J. Trop. Med. Hyg. 103 (4) (2020) 1621.

[7] I. Khaldarova, M. Pantti, 'Fake News: the Narrative Battle over the Ukrainian Conflict': 'The Future of Journalism: Risks, Threats and Opportunities', Routledge, 2020, pp. 228–238.
[8] S. van der Linden, J. Roozenbeek, 'Psychological inoculation against fake news', in: R. Greifeneder, M.E. Jaffé, E.J. Newman, N. Schwarz (Eds.), 'The Psychology of Fake News: Accepting, Sharing, and Correcting misinformation.', Routledge/Taylor & Francis Group, 2021, pp. 147–169.
[9] F.A. Ozbay, B. Alatas, 'Fake news detection within online social media using supervised artificial intelligence algorithms', Phys. Stat. Mech. Appl. 540 (2020), 123174.
[10] A.A.A. Ahmed, A. Aljabouh, P.K. Donepudi, M.S. Choi, 'Detecting Fake News Using Machine Learning: A Systematic Literature Review', 2021 arXiv preprint arXiv:2102.04458.
[11] M. Chorać, K. Demestichas, A. Giełczyk, Á. Herrero, P. Ksieniewicz, K. Remoundou, D. Urda, M. Woźniak, 'Advanced Machine Learning techniques for fake news (online disinformation) detection: a systematic mapping study', Appl. Soft Comput. 101 (2021), 107050.
[12] M.-L. Rusu, R.-E. Herman, 'Legislative measures adopted at the international level against fake news', in: 'Book Legislative Measures Adopted at the International Level against Fake News', 2 edn., Sciendo, 2019.
[13] S. Lewandowsky, S. van der Linden, 'Countering Misinformation and Fake News through Inoculation and Prebunking', European Review of Social Psychology, 2021.
[14] J. Roozenbeek, S. van der Linden, 'The fake news game: actively inoculating against the risk of misinformation', J. Risk Res. 22 (5) (2019) 570–580.
[15] L. Bode, E.K. Vraga, 'See Something, Say Something: Correction of Global Health Misinformation on Social Media', Health Communication, 2018, p. 1131.
[16] M. Tully, L. Bode, E.K. Vraga, 'Mobilizing Users: Does Exposure to Misinformation and its Correction Affect Users' Responses to a Health Misinformation Post?', vol. 6, Social Media + Society, 2020.
[17] E.K. Vraga, L. Bode, 'Correction as a solution for health misinformation on social media', Am. J. Publ. Health 110 (2020) S278–S280.
[18] L. Bode, E.K. Vraga, M. Tully, 'Correcting misperceptions about genetically modified food on social media: examining the impact of experts, social media heuristics, and the gateway belief model', Sci. Commun. 43 (2) (2021) 225–251.
[19] E.K. Vraga, L. Bode, 'Using expert sources to correct health misinformation in social media', Sci. Commun. 39 (5) (2017) 621–645.
[20] M. Hameleers, T.G.L.A. Van der Meer, 'Misinformation and polarization in a high-choice media environment: how effective are political fact-checkers?', Commun. Res. 47 (2) (2020) 227–250.
[21] E.K. Vraga, L. Bode, 'I do not believe you: how providing a source corrects health misperceptions across social media platforms', Inf. Commun. Soc. 21 (10) (2018) 1337–1353.
[22] N. Walter, S.T. Murphy, 'How to unring the bell: a meta-analytic approach to correction of misinformation', Commun. Monogr. 85 (3) (2018) 423–441.
[23] A. Chadwick, C. Vaccari, 'News sharing on UK social media: misinformation, disinformation, and correction', in: 'Book News Sharing on UK Social Media: Misinformation, Disinformation, and Correction', edn., 2019.
[24] E.C. Tandoc, D. Lim, R. Ling, 'Diffusion of disinformation: how social media users respond to fake news and why', Journalism 21 (3) (2020) 381–398.
[25] D.O. Vicol, 'Who is most likely to believe and to share misinformation? Full Fact', in: 'Book Who Is Most Likely to Believe and to Share Misinformation? Full Fact', edn., 2020.
[26] L. Bode, E.K. Vraga, 'Correction Experiences on Social Media during COVID-19', vol. 7, 2021. Social Media + Society.
[27] L. Bode, E.K. Vraga, 'Value for correction: documenting perceptions about peer correction of misinformation on social media in the context of COVID-19', Journal of Quantitative Description (2021) 1. Digital Media.
[28] J.K. Swim, L.L. Hyers, 'Excuse me—what did you just say?!: women's public and private responses to sexist remarks', J. Exp. Soc. Psychol. 35 (1) (1999) 68–88.
[29] J.K. Swim, L.L. Hyers, L.L. Cohen, D.C. Fitzgerald, W.H. Bylsma, 'African American college students' experiences with everyday racism: characteristics of and responses to these incidents', J. Black Psychol. 29 (1) (2003) 38–67.
[30] C.L. Dickter, V.A. Newton, 'To confront or not to confront: non-targets' evaluations of and responses to racist comments', J. Appl. Soc. Psychol. 43 (2013) E262–E275.
[31] S. Gurgun, E. Arden-Close, K. Phalp, R. Ali, 'Online Silence: Why Do People Not Challenge Others when Posting Misinformation?', Internet Research, 2022 (ahead-of-print).
[32] K. Chadha, L. Steiner, J. Vitak, Z. Ashktorab, 'Women's responses to online harassment', Int. J. Commun. 14 (2020) 19.
[33] J. Posetti, N. Aboulez, K. Bontcheva, J. Harrison, S. Waisbord, Online violence against women journalists, in: 'Book Online Violence against Women Journalists', edn., UNESCO, 2020.
[34] V.L. Byrne, "You might as well just all agree with each other:" an initial study of cyberbullying victims' social presence in online discussions', Comput. Educ. 167 (2021), 104174.
[35] R. Urbaniak, M. Ptaszyński, P. Tempska, G. Leliwa, M. Brochocki, M. Wroczyński, 'Personal attacks decrease user activity in social networking platforms', Comput. Hum. Behav. 126 (2022), 106972.
[36] E. Powers, M. Koliska, P. Guha, "Shouting matches and echo chambers": perceived identity threats and political self-censorship on social media', Int. J. Commun. 13 (2019) 20.
[37] A.F. Hayes, 'Exploring the forms of self-censorship: on the spiral of silence and the use of opinion expression avoidance strategies', J. Commun. (4) (2007) 785.
[38] A.Z.-X. Koo, M.-H. Su, S. Lee, S.-Y. Ahn, H. Rojas, 'What motivates people to correct misinformation? Examining the effects of third-person perceptions and perceived norms', J. Broadcast. Electron. Media 65 (1) (2021) 111–134.
[39] Y. Sun, S.C. Chia, F. Lu, J. Oktavianus, 'The battle is on: factors that motivate people to combat anti-vaccine misinformation', Health Commun. 37 (3) (2022) 327–336.
[40] F.J. Milliken, E.W. Morrison, P.F. Hewlin, 'An exploratory study of employee silence: issues that employees don't communicate upward and why', J. Manag. Stud. 40 (6) (2003) 1453–1476.
[41] L.V. Dyne, S. Ang, I.C. Botero, 'Conceptualizing employee silence and employee voice as multidimensional constructs', J. Manag. Stud. 40 (6) (2003) 1359–1392.
[42] N. Bienefeld, G. Grote, 'Silence that may kill', Aviat Psychol. Appl. Hum. Factors 2 (1) (2012) 1–10.
[43] J.R. Detert, Amy C. Edmondson, 'Implicit voice theories: Taken-for-granted rules of self-censorship at work' 54 (3) (2011) 461–488.
[44] A. Jaworski, I. Sachdev, 'Beliefs about silence in the classroom', Lang. Educ. 12 (4) (1998) 273–292.
[45] K.A. Rocca, 'Student participation in the college classroom: an extended multidisciplinary literature review', Commun. Educ. 59 (2) (2010) 185–213.
[46] P.A. Fassinger, 'Understanding classroom interaction: students' and professors' contributions to students' silence', J. High Educ. 66 (1) (1995) 82–96.
[47] S. Dingli, T.N. Cooke, 'Political Silence', Routledge, 2019, 2019.
[48] C.C. Pinder, K.P. Harlos, 'Employee Silence: Quiescence and Acquiescence as Responses to Perceived Injustice': 'Research in Personnel and Human Resources Management', Emerald Group Publishing Limited, 2001.
[49] M. Knoll, R. Van Dick, 'Do I hear the whistle? A first attempt to measure four forms of employee silence and their correlates', J. Bus. Ethics 113 (2) (2013) 349–362.
[50] N. Lapidot-Lefler, A. Barak, 'Effects of anonymity, invisibility, and lack of eye-contact on toxic online disinhibition', Comput. Hum. Behav. 28 (2) (2012) 434–443.
[51] J. Suler, 'The online disinhibition effect', Cyberpsychol. Behav. (3) (2004) 321.
[52] C.J. Ferguson, 'Does media violence predict societal violence? It depends on what you look at and when', J. Commun. 65 (1) (2015) E1–E22.
[53] J. Preece, B. Nonnecke, D. Andrews, 'The top five reasons for lurking: improving community experiences for everyone', Comput. Hum. Behav. 20 (2) (2004) 201–223.
[54] B. Nonnecke, J. Preece, 'Why Lurkers Lurk', 2001.
[55] J.W. Penney, 'Chilling effects: online surveillance and Wikipedia use', Berkeley Tech. LJ 31 (2016) 117.
[56] B. Marder, A. Joinson, A. Shankar, D. Houghton, 'The extended 'chilling' effect of Facebook: the cold reality of ubiquitous social networking', Comput. Hum. Behav. 60 (2016) 582–592.
[57] J.M. Hudson, A.S. Bruckman, 'The bystander effect: a lens for understanding patterns of participation', J. Learn. Sci. 13 (2) (2004) 165–195.
[58] M.F. Beaudoin, 'Learning or lurking?: tracking the "invisible" online student', Internet High Educ. 5 (2) (2002) 147–155.
[59] N.J. Stroud, E. Van Duyn, C. Peacock, 'News Commenters and News Comment Readers', Engaging News Project, 2016, pp. 1–21.
[60] K.N. Hampton, H. Rainie, W. Lu, M. Dwyer, I. Shin, K. Purcell, 'Social media and the'spiral of silence'', in: 'Book Social Media and The'spiral of Silence'' (Pew Research Center, edn., 2014.
[61] C. McClain, American trend panel, in: 'Book American Trend Panel', edn., 2021.
[62] E. Noelle-Neumann, 'The Spiral of Silence: Public Opinion–Our Social Skin', University of Chicago Press, 1993, 1993.
[63] E. Noelle-Neumann, 'The spiral of silence a theory of public opinion', J. Commun. 24 (2) (1974) 43–51.
[64] A.D. Berkowitz, 'An overview of the social norms approach', Changing the culture of college drinking, A socially situated health communication campaign 1 (2005) 193–214.
[65] S. Moscovici, G. Mugny, 'Minority Influence': 'Basic Group Processes', Springer, 1983, pp. 41–64.
[66] W. Wood, S. Lundgren, J.A. Ouellette, S. Busceme, T. Blackstone, 'Minority influence: a meta-analytic review of social influence processes', Psychol. Bull. 115 (3) (1994) 323.
[67] J.W. Bolderdijk, L. Jans, 'Minority influence in climate change mitigation', Current Opinion in Psychology 42 (2021) 25–30.
[68] B. Ross, L. Pilz, B. Cabrera, F. Brachten, G. Neubaum, S. Stieglitz, 'Are social bots a real threat? An agent-based model of the spiral of silence to analyse the impact of manipulative actors in social networks', Eur. J. Inf. Syst. 28 (4) (2019) 394–412.
[69] C. Carter Olson, V. LaPoe, 'Combating the Digital Spiral of Silence: Academic Activists versus Social Media Trolls', Mediating Misogyny: Gender, Technology, and harassment, 2018, pp. 271–291.
[70] E. Pariser, 'The Filter Bubble: what the Internet Is Hiding from You', Penguin UK, 2011, 2011.
[71] B. Semaan, H. Faucett, S.P. Robertson, M. Maruyama, S. Douglas, 'Designing political deliberation environments to support interactions in the public sphere', in: 'Book Designing Political Deliberation Environments to Support Interactions in the Public Sphere', edn., 2015, pp. 3167–3176.
[72] C. McLaughlin, J. Vitak, 'Norm evolution and violation on Facebook', New Media Soc. 14 (2) (2012) 299–315.
[73] E.K. Vraga, K. Thorson, N. Kligler-Vilenchik, E. Gee, 'How individual sensitivities to disagreement shape youth political expression on Facebook', Comput. Hum. Behav. 45 (2015) 281–289.

[74] M. Ekström, 'Young people's everyday political talk: a social achievement of democratic engagement', J. Youth Stud. 19 (1) (2016) 1–19.
[75] K.W. Thomas, 'Conflict and conflict management: reflections and update', J. Organ. Behav. (1992) 265–274.
[76] H. Lee, 'Behavioral strategies for dealing with flaming in an online forum', Socio. Q. 46 (2) (2005) 385–403.
[77] D. Tjosvold, H.F. Sun, 'Understanding Conflict Avoidance: Relationship, Motivations, Actions, and Consequences', International Journal of Conflict Management, 2002.
[78] C. Tagg, P. Seargeant, A.A. Brown, 'Taking offence on social media : conviviality and communication on Facebook', in: 'Book Taking Offence on Social Media : Conviviality and Communication on Facebook', Palgrave Macmillan, 2017 imprint published by Springer Nature.
[79] M. Alonzo, M. Aiken, 'Flaming in electronic communication', Decis. Support Syst. 36 (3) (2004) 205–213.
[80] C. Tagg, P. Seargeant, 'Negotiating Social Roles in Semi-public Online Contexts', Social Media Discourse,(dis) Identifications and Diversities, 2016.
[81] A. Chadwick, C. Vaccari, N.-A. Hall, 'Covid vaccines and online personal messaging: the challenge of challenging everyday misinformation', in: 'Book Covid Vaccines and Online Personal Messaging: the Challenge of Challenging Everyday Misinformation', edn., Loughborough University, 2022.
[82] J.R. Bautista, Y. Zhang, J. Gwizdka, 'US physicians' and nurses' motivations, barriers, and recommendations for correcting health misinformation on social media: qualitative interview study', JMIR Public Health and Surveillance 7 (9) (2021), e27715.
[83] S.F. Waterloo, S.E. Baumgartner, J. Peter, P.M. Valkenburg, 'Norms of online expressions of emotion: comparing Facebook, Twitter, Instagram, and WhatsApp', New media & society 20 (5) (2018) 1813–1831.
[84] I.V. Pasquetto, E. Jahani, S. Atreja, M. Baum, 'Social debunking of misinformation on WhatsApp: the case for strong and in-group ties', Proceedings of the ACM on Human-Computer Interaction 6 (CSCW1) (2022) 1–35.
[85] Q. Statista, 'Number of monthly active Facebook users worldwide as of 4th quarter 2015', in: 'Book Number of Monthly Active Facebook Users Worldwide as of 4th Quarter 2015', edn., 2016.
[86] https://www.statista.com/statistics/376128/facebook-global-user-age-distribution/, 2022. March 2022.
[87] I. Centre for International Governance, Ipsos: 'CIGI-Ipsos Global Survey on Internet Security and Trust', 2019.
[88] N.B. Ellison, D. Boyd, 'Sociality through Social Network Sites', The Oxford handbook of internet studies, 2013, pp. 151–172.
[89] P. De Meo, E. Ferrara, G. Fiumara, A. Provetti, 'On Facebook, most ties are weak', Commun. ACM 57 (11) (2014) 78–84.
[90] M.S. Granovetter, 'The strength of weak ties', Am. J. Sociol. 78 (6) (1973) 1360–1380.
[91] Y. Kim, 'The contribution of social network sites to exposure to political difference: the relationships among SNSs, online political messaging, and exposure to cross-cutting perspectives', Comput. Hum. Behav. 27 (2) (2011) 971–977.
[92] S. Valenzuela, Y. Kim, H. Gil de Zúñiga, 'Social networks that matter: exploring the role of political discussion for online political participation', Int. J. Publ. Opin. Res. 24 (2) (2012) 163–184.
[93] E.L. Cohen, A. Atwell Seate, S.M. Kromka, A. Sutherland, M. Thomas, K. Skerda, A. Nicholson, 'To correct or not to correct? Social identity threats increase willingness to denounce fake news through presumed media influence and hostile media perceptions', Commun. Res. Rep. 37 (5) (2020) 263–275.
[94] S. Tangirala, R. Ramanujam, 'Employee silence on critical work issues: the cross level effects of procedural justice climate', Person. Psychol. 61 (1) (2008) 37–68.
[95] A. Arif, J.J. Robinson, S.A. Stanek, E.S. Fichet, P. Townsend, Z. Worku, K. Starbird, 'A closer look at the self-correcting crowd: examining corrections in online rumors'Proc. Proceedings of the, in: ACM Conference on Computer Supported Cooperative Work and Social Computing, Portland, Oregon, 2017, p. USA2017.
[96] S. Gearhart, W. Zhang, ''Was it something I said?' 'No, it was something you posted!' A study of the spiral of silence theory in social media contexts', Cyberpsychol., Behav. Soc. Netw. 18 (4) (2015) 208–213.
[97] G. Neubaum, N.C. Krämer, 'What do we fear? Expected sanctions for expressing minority opinions in offline and online communication', Commun. Res. 45 (2) (2018) 139–164.
[98] Y. Theocharis, A. Cardenal, S. Jin, T. Aalberg, D.N. Hopmann, J. Strömbäck, L. Castro, F. Esser, P. Van Aelst, C. de Vreese, 'Does the Platform Matter? Social Media and COVID-19 Conspiracy Theory Beliefs in 17 Countries', new media & society, 2021, 14614448211045666.
[99] C. Geeng, S. Yee, F. Roesner, 'Fake News on Facebook and Twitter: Investigating How People (Don't) Investigate', Conference on Human Factors in Computing Systems Proceedings, 2020, pp. 1–14.
[100] P. Moy, D. Domke, K. Stamm, 'The spiral of silence and public opinion on affirmative action', Journal. Mass Commun. Q. 78 (1) (2001) 7–25.
[101] S. Gearhart, W. Zhang, 'Gay bullying and online opinion expression', Soc. Sci. Comput. Rev. 32 (1) (2014) 18–36.
[102] F. Faul, E. Erdfelder, A.-G. Lang, A. Buchner, 'G* Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences', Behav. Res. Methods 39 (2) (2007) 175–191.
[103] S.B. Green, 'How many subjects does it take to do a regression analysis', Multivariate Behav. Res. 26 (3) (1991) 499–510.
[104] F.D. Schönbrodt, M. Perugini, 'At what sample size do correlations stabilize?', J. Res. Pers. 47 (5) (2013) 609–612.
[105] O.Y. Khasawneh, 'Technophobia: examining its hidden factors and defining it', Technol. Soc. 54 (2018) 93–100.
[106] K. Rouibah, H. Abbas, S. Rouibah, 'Factors affecting camera mobile phone adoption before e-shopping in the Arab world', Technol. Soc. 33 (3–4) (2011) 271–283.
[107] M.C. Howard, 'A review of exploratory factor analysis decisions and overview of current practices: what we are doing and how can we improve?', Int. J. Hum. Comput. Interact. 32 (1) (2016) 51–62.
[108] H.F. Kaiser, 'An index of factorial simplicity', Psychometrika 39 (1) (1974) 31–36.
[109] R.B. Cattell, 'The scree test for the number of factors', Multivariate Behav. Res. 1 (2) (1966) 245–276.
[110] C. DiStefano, M. Zhu, D. Mindrila, 'Understanding and using factor scores: considerations for the applied researcher', Practical Assess. Res. Eval. 14 (1) (2009) 20.
[111] C.D. Anderson, L.B. Stephenson, 'Mobilizing the young: the role of social networks', Canadian Journal of Political Science/Revue canadienne de science politique 51 (4) (2018) 861–880.
[112] M.X. Delli Carpini, 'Gen.com: youth, civic engagement, and the new information environment', Polit. Commun. 17 (4) (2000) 341–349.
[113] K. Thorson, 'Facing an uncertain reception: young citizens and political interaction on Facebook', Inf. Commun. Soc. 17 (2) (2014) 203–216.
[114] M. Department for Culture, Sport: 'Taking Part Focus on: Social Media', 2016.
[115] C.F. Scott, L.Y. Bay-Cheng, M.A. Prince, T.H. Nochajski, R.L. Collins, 'Time spent online: latent profile analyses of emerging adults' social media use', Comput. Hum. Behav. 75 (2017) 311–319.
[116] F.T. McAndrew, H.S. Jeong, 'Who does what on Facebook? Age, sex, and relationship status as predictors of Facebook use', Comput. Hum. Behav. 28 (6) (2012) 2359–2365.
[117] J.D. Henry, W. von Hippel, M.R. Nangle, M. Waters, 'Age and the experience of strong self-conscious emotion', Aging Ment. Health 22 (4) (2018) 497–502.
[118] U. Orth, R.Y. Erol, E.C. Luciano, 'Development of self-esteem from age 4 to 94 years: a meta-analysis of longitudinal studies', Psychol. Bull. 144 (10) (2018) 1045–1080.
[119] E.C. Tandoc Jr., H.K. Kim, 'Avoiding real news, believing in fake news? Investigating pathways from information overload to misbelief', Journalism 24 (6) (2023) 1174–1192.
[120] M. Goyanes, A. Ardèvol-Abreu, H. Gil de Zúñiga, 'Antecedents of news avoidance: competing effects of political interest, news overload, trust in news media, and "news finds me" perception', Digital Journalism 11 (1) (2023) 1–18.
[121] C. Peacock, E. Van Duyn, 'Monitoring and Correcting: Why Women Read and Men Comment Online', Information, Communication & Society, 2021, pp. 1–16.
[122] D. Rosenberg, R. Mano, G.S. Mesch, 'Absolute monopoly, areas of control or democracy? Examining gender differences in health participation on social media', Comput. Hum. Behav. 102 (2020) 166–171.
[123] L. Bode, 'Closing the gap: gender parity in political engagement on social media', Inf. Commun. Soc. 20 (4) (2017) 587–603.
[124] X. Chen, S.-C.J. Sin, Y.-L. Theng, C.S. Lee, 'Why do social media users share misinformation?', in: 'Book Why Do Social Media Users Share Misinformation?', edn., 2015, pp. 111–114.
[125] S. Altay, A.-S. Hacquin, H. Mercier, 'Why Do So Few People Share Fake News? it Hurts Their Reputation', New Media Society, 2019, 1461444820969893.
[126] D. Lazer, M. Baum, N. Grinberg, L. Friedland, K. Joseph, W. Hobbs, C. Mattsson, 'Combating Fake News: an Agenda for Research and Action', 2017.
[127] T. Hopp, 'Fake news self-efficacy, fake news identification, and content sharing on Facebook', J. Inf. Technol. Polit. 19 (2) (2022) 229–252.
[128] A. Bandura, 'Self-efficacy : the Exercise of Control', W.H. Freeman and Company, 1997, 1997.
[129] H. Tajfel, J.C. Turner, 'The Social Identity Theory of Intergroup Behavior': 'Political Psychology', Psychology Press, 2004, pp. 276–293.
[130] A. Whiting, D. Williams, 'Why people use social media: a uses and gratifications approach', Qualitative Market Research (4) (2013) 362.
[131] A.J.M. Wagner, 'Do not click "like" when somebody has died: the role of norms for mourning practices in social media', Social Media+ Society 4 (1) (2018), 2056305117744392.
[132] D.L. Paddock, B.T. Bell, "It's better saying I look fat instead of saying you look fat": a Qualitative Study of UK Adolescents' Understanding of Appearance-Related Interactions on Social Media', J. Adolesc. Res. (2021), 07435584211034875.
[133] J.B. Sabra, "I hate when they do that!" Netiquette in mourning and memorialization among Danish Facebook users', J. Broadcast. Electron. Media 61 (1) (2017) 24–40.
[134] D. Geddes, R.A. Baron, 'Workplace aggression as a consequence of negative performance feedback', Manag. Commun. Q. 10 (4) (1997) 433–454.
[135] R.G. Bing-You, R.L. Trowbridge, 'Why medical educators may be failing at feedback', JAMA 302 (12) (2009) 1330–1331.
[136] C.H. Jeffries, M.J. Hornsey, 'Withholding negative feedback: is it about protecting the self or protecting others?', Br. J. Soc. Psychol. 51 (4) (2012) 772–780.
[137] A.M.B. Gallrein, K.L. Bollich-Ziegler, D. Leising, 'Interpersonal feedback in everyday life: empirical studies in Germany and the United States', Eur. J. Soc. Psychol. 49 (1) (2019) 1–18.
[138] S. Das, A. Kramer, 'Self-censorship on facebook', in: 'Book Self-Censorship on Facebook', 1 edn., 2013, pp. 120–127.
[139] P.B. Brandtzæg, J. Heim, 'Why people use social networking sites', in: 'Book Why People Use Social Networking Sites', edn., Springer, 2009, pp. 143–152.
[140] A. Nadkarni, S.G. Hofmann, 'Why do people use Facebook?', Pers. Indiv. Differ. 52 (3) (2012) 243–249.

[141] A.E. Marwick, D. Boyd, 'I tweet honestly, I tweet passionately: Twitter users, context collapse, and the imagined audience', New Media Soc. 13 (1) (2011) 114–133.
[142] S. Rosen, A. Tesser, 'On Reluctance to Communicate Undesirable Information: the MUM Effect', Sociometry, 1970, pp. 253–263.
[143] A.L. Gonzales, J.T. Hancock, 'Mirror, mirror on my Facebook wall: effects of exposure to Facebook on self-esteem', Cyberpsychol., Behav. Soc. Netw. 14 (1–2) (2011) 79–83.
[144] J. Kim, J. Ahn, 'The show must go on: the presentation of self during interpersonal conflict on Facebook', Proceedings of the American Society for Information Science and Technology 50 (1) (2013) 1–10.
[145] Z.-X. Zhang, X. Wei, 'Superficial harmony and conflict avoidance resulting from negative anticipation in the workplace', Manag. Organ. Rev. 13 (4) (2017) 795–820.
[146] R.W. Rogers, S. Prentice-Dunn, 'Protection Motivation Theory', 1997.
[147] C.D. Batson, J.G. Batson, J.K. Slingsby, K.L. Harrell, H.M. Peekna, R.M. Todd, 'Empathic joy and the empathy-altruism hypothesis', J. Pers. Soc. Psychol. 61 (3) (1991) 413.
[148] C.D. Batson, D.A. Lishner, E.L. Stocks, 'The Empathy—Altruism Hypothesis', 2015.
[149] R.E. Petty, J.T. Cacioppo, 'Issue involvement can increase or decrease persuasion by enhancing message-relevant cognitive responses', J. Pers. Soc. Psychol. 37 (10) (1979) 1915.
[150] A.G. Yong, S. Pearce, 'A beginner's guide to factor analysis: focusing on exploratory factor analysis', Tutorials in quantitative methods for psychology 9 (2) (2013) 79–94.
[151] V. Venkatesh, M.G. Morris, P.L. Ackerman, 'A longitudinal field investigation of gender differences in individual technology adoption decision-making processes', Organ. Behav. Hum. Decis. Process. 83 (1) (2000) 33–60.
[152] V. Venkatesh, M.G. Morris, 'Why don't men ever stop to ask for directions?. Gender, Social Influence, and Their Role in Technology Acceptance and Usage Behavior' MIS quarterly, 2000, pp. 115–139.
[153] Z. Zhou, X.-L. Jin, Y. Fang, 'Moderating role of gender in the relationships between perceived benefits and satisfaction in social virtual world continuance', Decis. Support Syst. 65 (2014) 69–79.
[154] H.A. Butler, 'Halpern Critical Thinking Assessment predicts real-world outcomes of critical thinking', Appl. Cognit. Psychol. 26 (5) (2012) 721–729.
[155] H.A. Butler, C.P. Dwyer, M.J. Hogan, A. Franco, S.F. Rivas, C. Saiz, L.S. Almeida, 'The halpern critical thinking assessment and real-world outcomes: cross-national applications', Think. Skills Creativ. 7 (2) (2012) 112–121.
[156] M.S. Bernstein, E. Bakshy, M. Burke, B. Karrer, 'Quantifying the invisible audience in social networks', in: 'Book Quantifying the Invisible Audience in Social Networks', edn., 2013, pp. 21–30.
[157] M. Sleeper, R. Balebako, S. Das, A.L. McConahy, J. Wiese, L.F. Cranor, 'The post that wasn't: exploring self-censorship on facebook'. Proc. Proceedings of the 2013 Conference on Computer Supported Cooperative Work, San Antonio, Texas, 2013, p. USA.
[158] D. Choi, S. Chun, H. Oh, J. Han, T. Kwon, 'Rumor propagation is amplified by echo chambers in social media', Sci. Rep. 10 (1) (2020) 1–10.
[159] M. Del Vicario, A. Bessi, F. Zollo, F. Petroni, A. Scala, G. Caldarelli, H.E. Stanley, W. Quattrociocchi, 'The spreading of misinformation online', Proc. Natl. Acad. Sci. USA 113 (3) (2016) 554–559.
[160] C. von der Weth, J. Vachery, M. Kankanhalli, 'Nudging users to slow down the spread of fake news in social media', in: 'Book Nudging Users to Slow Down the Spread of Fake News in Social Media', edn., IEEE, 2020, pp. 1–6.
[161] M. Tully, E.K. Vraga, L. Bode, 'Designing and testing news literacy messages for social media', Mass Commun. Soc. 23 (1) (2020) 22–46.
[162] T. Krauthoff, M. Baurmann, G. Betz, M. Mauve, 'Dialog-Based online argumentation', in: 'Book Dialog-Based Online Argumentation', edn., 2016, pp. 33–40.
[163] C. Meter, T. Krauthoff, M. Mauve, 'discuss: embedding dialog-based discussions into websites', in: 'Book Discuss: Embedding Dialog-Based Discussions into Websites', edn., Springer, 2017, pp. 449–460.
[164] F. Gao, C.X. Wang, Y. Sun, 'New model of productive online discussion and its implications for research and instruction', Journal of Educational Technology Development and Exchange 2 (1) (2009) 65.
[165] S. Sudman, N.M. Bradburn, N. Schwarz, 'Thinking about Answers: the Application of Cognitive Processes to Survey Methodology', Jossey-Bass, 1996, 1996.
[166] P. Grimm, 'Social Desirability Bias', Wiley international encyclopedia of marketing, 2010.