ARTICLE

Check for updates

Designing social media to foster user engagement in challenging misinformation: a cross-cultural comparison between the UK and Arab countries

Muaadh Noman[1], Selin Gurgun[2], Keith Phalp[2] & Raian Ali [1✉]

Challenging others who post misinformation is a type of social correction that complements algorithm-based approaches. However, participation rates in such social acts remain limited. In this paper, we study design techniques that leverage principles of persuasive system design and communication theories to foster such prosocial behaviour across two distinct cultural contexts: the British and the Arab. A total of 462 participants completed an online survey (250 UK, 212 Arabs). The study compared the two cultural contexts regarding willingness to challenge misinformation and the persuasiveness of seven design techniques to increase that willingness, namely predefined question stickers, thinking face reaction, sentence openers, fact checker badge, social norm messages, tone detector, and private commenting. Moreover, it explores the impact of individuals' characteristics on their perception of the techniques as being more or less persuasive than a standard comment box. The study found that the willingness to challenge misinformation was significantly higher in the Arab context than in the UK context. Moreover, except for the private commenting, all techniques were more impactful in the Arab context than in the UK context. Some techniques, such as predefined question stickers, were more effective in both cultures compared to the standard comment box, while others, like the fact checker badge, were more effective only in the Arab context. However, in the UK, sentence openers had a lower impact. Furthermore, personality traits, age, and perspective-taking showed the potential but also the varying impacts on the persuasiveness of the techniques on users' correction of misinformation across both cultural contexts while pointing to the need for considering both personal and cultural factors in designing social-correction-based solutions.

[1] College of Science and Engineering, Hamad Bin Khalifa University, Doha, Qatar. [2] Faculty of Science and Technology, Bournemouth University, Poole, UK. ✉email: raali2@hbku.edu.qa

Introduction

The rapid spread of misinformation has recently received much attention (Aïmeur et al. 2023). While the issue of misinformation is not new, its impact extends across multiple domains, ranging from political situations (Kuklinski et al. 2000) to health-related issues (Lewandowsky et al. 2012). The popularity of social media platforms has significantly exacerbated this situation by allowing for rapid transmission and extensive sharing across varied audiences (Vicario et al. 2016; Vosoughi et al. 2018). Long-term exposure to misinformation has been shown to increase the likelihood of individuals accepting its falsehoods (Lewandowsky et al. 2012; Pan et al. 2021), a phenomenon known as the "truth by repetition" effect, in which repeated statements are given more credibility than novel assertions (Morgan and Cappella, 2023). Anchoring, another cognitive bias where an individual heavily relies on an initial piece of information to make subsequent judgements (Tversky and Kahneman, 1974), can further exacerbate the impact of misinformation by influencing people to remain attached to false information, even when presented with credible countering evidence, thus impacting decision-making in a significant way (Jost et al. 2020). As a result, it is critical to develop effective tactics for countering misinformation to avoid any possible harm (Bode and Vraga, 2015; Nyhan and Reifler, 2010).

Various approaches have been suggested to combat misinformation. These include inoculating individuals against misinformation (Cook et al. 2017), automating the process of detecting and correcting it (Choraś et al. 2021; Vicari and Komendatova, 2023), and enabling social media users to rectify false information (Bode and Vraga, 2017; Kligler-Vilenchik, 2022). Social corrections have shown promise in mitigating the spread of misinformation (Walter and Murphy, 2018). Furthermore, information interventions, including those offering new or corrective data for misinformation, play a crucial role in shaping perceptions of the content on social media platforms (Bao et al. 2022). However, research indicates that users frequently show reticence towards addressing misinformation when encountering it on social media (Chadwick et al. 2023; Tandoc Jr et al. (2020)). Gurgun et al. (2022) identified six types of reasons that may influence users' willingness to challenge individuals posting misinformation: technical aspects, content-oriented, self-oriented, others-oriented, individual characteristics, and relationship-oriented. Substantial research, underpinned by cognitive psychology (Eysenck and Keane, 2015), has been undertaken to tackle these obstacles, providing insight into individuals' reactions to misinformation and their endeavours to correct it. For instance, Lewandowsky et al. (2012) explored how cognitive psychology theories can be leveraged by technology-based solutions for creating effective interventions. While these approaches show promise in combating misinformation, it is crucial to consider the design of online social media platforms to effectively address this issue.

Persuasive design techniques incorporated into technology-based solutions have been used to influence behaviour change (Oinas-Kukkonen, 2013). Oinas-Kukkonen and Harjumaa (2009) proposed the Persuasive System Design (PSD) model, which is a structured framework that systematically categorises the persuasive principles in system design into four categories: Primary Task Support, Dialogue Support, System Credibility Support, and Social Support. Each category consists of seven specific persuasive principles to influence users' behaviours and attitudes. The Primary Task Support category promotes ease of use and user engagement. It includes principles such as Reduction, which emphasises reducing the complexity and steps in doing tasks; Tunnelling, which guides users through a process or experience; and Self-monitoring, which emphasises allowing system users to monitor and adjust their performance. The Dialogue Support category fosters a more personalised interaction. It consists of principles such as using visually attractive design elements, known as Liking, and offering fitting suggestions to users, known as Suggestion. The System Credibility Support category focuses on building trust. It includes principles such as Verifiability, which is demonstrated by providing means for content accuracy verification through outside resources. Finally, the Social Support category contains principles that leverage social dynamics to motivate attitudinal and behavioural change, such as Normative Influence, demonstrated by displaying norms to foster people's behaviour, and Recognition, which emphasises publicly recognising individuals exhibiting specific behaviours. PSD has been extensively used as a reference model in technology-assisted behaviour change solutions, e.g., in the domain of digital health (McGowan et al. 2022) and combating gambling problems (Cemiloglu et al. 2023). Studies have demonstrated that persuasive design techniques can effectively encourage people to change their behaviour in a variety of contexts, including social media (Elaheebocus et al. 2018; Wiafe et al. 2020), so that these techniques could also be used to encourage users to confront misinformation on online platforms.

One of the few studies that has addressed behaviours involving interactions with others in the context of challenging misinformation is the study by Gurgun et al. (2023). This study draws upon the PSD model to create social media interfaces that include persuasive design techniques to encourage users to confront misinformation. Gurgun et al. (2023) suggested seven design techniques that were modelled after Facebook's current interface. The first technique was "Private Commenting," which allows users to send private comments directly on misinformation posts. The second technique, "Predefined Question Stickers," provides pre-written questions for users when challenging information. The third technique was "Tone Detector," which indicates the emotional tone of a user's comments. The fourth technique is the "Fact Checker Badge," awarded for correcting misinformation. The fifth technique was "Social Norm Messages," highlighting the community's positive stance on misinformation correction through pop-up messages. The sixth technique was "Sentence Openers," which assists users in countering misinformation by suggesting the first part of the argument. Lastly, the "Thinking Face Reaction" emphasises using visually attractive reactions to express scepticism about the misinformation. Cross-cultural studies on online behaviours, including challenging misinformation, have mostly used WEIRD (western, educated, industrialised, rich, and democratic) samples (Henrich et al. 2010), which potentially limits the applicability of their findings to non-WEIRD cultures like Arab countries. Gurgun et al. (2023) study was conducted within the cultural context of the United Kingdom (UK). Therefore, further research is needed to explore how individual and cultural factors influence the effectiveness of persuasive techniques in motivating users to challenge misinformation in other cultural contexts, such as those in Arab countries.

The impact of persuasive design techniques on behaviour and decision-making may vary depending on individual factors such as personality traits, empathy, perspective-taking, and age. Personality traits have been shown to influence the willingness to participate in discussions. For instance, extroverts are typically more willing to engage in interactions (Blau and Barak, 2012). In addition, the literature suggests a significant relationship between persuasive strategies and personality traits (Alqahtani et al. 2022; Halko and Kientz, 2010). According to Gurgun et al. (2023), increasing openness to new experiences is associated with an increased likelihood of the perceived persuasiveness of design

techniques like sending a thinking face reaction to someone who posts misinformation. Moreover, Gurgun et al. (2023) have examined the associations between age, perspective-taking, empathic concern, and persuasiveness of a range of design techniques that social media can employ to motivate users to challenge those who post misinformation. The study showed that increasing age and perspective-taking were associated with an increased perception of the persuasiveness of the design techniques such as "predefined question stickers." We argue that the relationship between personal factors and the perceived persuasiveness of the techniques is influenced by the larger cultural context of the user. This is consistent with the prior research in cross-cultural human-computer interaction (HCI), which highlights the importance of cultural factors for designing effective user-centric interfaces (Adnan et al. 2020; Ford and Kotzé, 2005; Reinecke and Bernstein, 2011).

Culture refers to the distinct patterns of thinking and behaviour shared within a specific group (Hofstede, 2001). It encompasses various aspects, including communication styles, which play a crucial role in shaping the interactions within a society, including those needed when challenging and correcting others. In the study of cultural distinctions, scholars have used different frameworks to analyse these variations, such as Edward Hall's high-low context theory (Hall, 1976). That context theory categorises societies into two groups based on their communication styles: high-context cultures and low-context cultures. In high-context cultures, such as those in Arab countries, communication relies less on verbal expressions and more on non-verbal cues derived from the situational context. People in these cultures often infer meaning from the surrounding environment, including body language, gestures, and social hierarchies. On the other hand, low-context cultures, like the UK, tend to have more formalised communication preferences, where individuals prefer to convey information through explicit verbal forms. Clear and direct communication is valued in these cultures, and the emphasis is placed on the precise exchange of information.

These differences in communication styles across cultures are expected to pose challenges to the way social media should be designed, as they are meant to be used by users from all over the world. The differences shall also affect the persuasiveness of the design techniques for challenging misinformation proposed by Gurgun et al. (2023), as they rank differently with respect to being closer to low- or high-context communication. Therefore, software designers may need to choose and implement design techniques that facilitate combating misinformation according to communication styles. This notion aligns with the "digital nudging" concept derived from behavioural economics, positing that subtle alterations in the digital environment, such as user interface designs, can influence user decisions (Hummel and Maedche, 2019; Weinmann et al. 2016). When implementing persuasive design techniques to combat misinformation in high-context cultures, attention may be given to the significance of non-verbal cues and contextual elements. Incorporating visual cues, symbols, and context-specific references into these interfaces can better resonate with the communication patterns of high-context cultures. Conversely, in low-context cultures where explicit verbal communication is valued, persuasive techniques may prioritise clear and concise messaging. By taking into account the preferences for direct communication, design interventions can be developed that appeal to individuals in low-context cultures and foster their active participation in combating misinformation. Hence, there is a need to examine whether the persuasiveness of the techniques proposed in Gurgun et al. (2023) varies across cultures, particularly in high-context cultures such as the Arab culture.

In this study, we aim to replicate and expand the research conducted by Gurgun et al. (2023) and to replicate it in a different cultural context, specifically in the Arab context. Our study aims to inform how persuasive design techniques can be tailored to align with the communication patterns of high-context cultures, considering the significance of non-verbal cues and contextual elements. We first examine the differences in the perceived persuasiveness of each of the techniques with regard to the behaviour of challenging misinformation, and then we study how they differ in comparison to the standard comment box, typically found on social media platforms. We then examine the role of demographics, personality traits, perspective-taking, and empathic concern in users' perceived persuasiveness of the techniques. The insights gained from this study can inform future interface design and interventions, enabling more tailored and culturally sensitive approaches to combat misinformation.

This study aims to answer the following research questions:

RQ1: Is there a difference in the persuasiveness of social media design techniques to challenge misinformation between UK and Arab users?

RQ2: Is there an impact of gender, age, empathy, perspective-taking, and personality traits on perception of design techniques as being more or less persuasive than a standard comment box among users from the UK and the Arab countries?

Research Design

Questionnaire Design. The questionnaire for this study was developed using the Qualtrics platform (https://www.qualtrics.com). Essential terminologies, such as "challenging" and "misinformation," were explicitly explained at the beginning of the questionnaire to ensure an unambiguous understanding of these terms by the participants. Our study was purposefully directed towards challenging acquaintances, denoting connections on social media that fall somewhere between strangers and close friends, including but not limited to previous co-workers, neighbours, or people belonging to the same social media groups. We needed to make this clarification because previous studies have suggested that people may behave differently when they interact with acquaintances versus strangers and also versus relatives and close friends. Specifically, people might hesitate to challenge misinformation posted by acquaintances because they worry about how it might affect their relationships with them (Valenzuela et al. 2012).

We explored the impact of seven design techniques (Table 1) that can be integrated into the design of social media on users' willingness to challenge acquaintances who post misinformation. These techniques were grounded in the Persuasive Systems Design (PSD) model (Oinas-Kukkonen and Harjumaa, 2009). They were modelled after Facebook's current interface. The participants were introduced to these design techniques through a scenario featuring a widely shared misinformation news article about a potential asteroid collision with the Earth, originally published on CNN's iReport news hub in 2014 (Matyszczyk, 2014). We informed the participants that the news in the scenario was false, as our study aimed to assess perceptions of the persuasiveness of the techniques, not discerning misinformation. We intentionally gave the Facebook account sharing the false news a gender-neutral name and photo in order to convey a non-anonymous identity. The participants were informed that this account belonged to an acquaintance, thus controlling for familiarity factors. Figure 1 shows an example of the "fact checker badge" prototype. The "fact checker badge" is intended to acknowledge and incentivise users through recognition. This badge enables users to express their agreement with comments that challenge misinformation. Individuals who receive a specific

Table 1 Persuasive design techniques (adapted from Gurgun et al. 2023).

Design Technique	Description
Private Commenting (PC)	It simplifies the task of challenging misinformation by allowing users to send direct private comments on posts instead of private message. It is developed based on the **reduction** principle, which emphasises reducing complexity and steps of doing tasks.
Predefined Question Stickers (PQS)	It offers preset questions that facilitate users to challenge, e.g., stickers with labels like "What is your source?". It is grounded in the strategy of offering fitting suggestions called **suggestion**.
Tone Detector (TD)	It provides real-time feedback on the emotional tone of users' comments (e.g., friendly, aggressive), allowing them to monitor and adjust their own responses. It is based on the **self-monitoring** principle.
Fact Checker Badge (FCB)	A badge is awarded to users who rectify misinformation. This design technique is based on the **recognition** principle, which emphasises publicly recognising individuals exhibiting specific behaviours.
Social Norm Message (SNM)	Pop-up prompts show social media users' acceptance of combating misinformation. It's grounded on the principle of displaying norms to foster people's behaviour, known as **normative influence**.
Sentence Openers (SO)	It offers preset sentence openers that facilitate users to challenge, e.g., "My argument is.." This is rooted in the **tunnelling** principle, which guides users through a process or experience.
Thinking Face Reaction (TFR)	It provides users with an expressive reaction to convey their doubts about the content they encounter. It's grounded on the principle of the persuasiveness of visually attractive design elements, known as **liking**.

Fig. 1 Fact Checker Badge high-fidelity prototype: in English (left) and Arabic (right). This figure illustrates a user's comment challenging misinformation (top) and the community's awarding of the Fact Checker Badge to the user's comment to recognise their engagement (bottom).

number of votes have the ability to display the "fact checker badge" on their profile. Consequently, it will be apparent to other users that the individual adorned with the badge has demonstrated proactive behaviour in combating misinformation disseminated on social media platforms.

The survey design, in both English and Arabic, is accessible on the Open Science Framework (OSF) (https://osf.io/cys8j/).

Participants and procedure. A total of 462 participants completed an online survey (250 UK, 212 Arabs). To participate in our study, UK participants had to meet the following criteria: be at least 18 years old, speak English fluently, have an active Facebook account with a real identity, encounter misinformation on Facebook, and live in the UK. Arab participants had to meet the same criteria, except that they had to speak Arabic fluently and live, not necessarily for the entirety of their lives, in an Arab country.

We used the Prolific (www.prolific.co) platform to recruit participants from both cultural contexts. We also used Cint (www.cint.com) to get responses from Arab participants. Participants who fulfilled the eligibility criteria consented to participate in an online survey. They were informed that they could withdraw at any time. The questionnaire contained three attention checks, and participants who did not pass two or more of these checks were subsequently disqualified from the study. We

also excluded participants who completed the survey within a very short time, provided contradictory answers, or exhibited response patterns such as straight-line or zigzag patterns while responding to the items measured on the Likert scale. The presence of gibberish text in response to open-ended questions was also considered an indicator of non-authentic responses. The eligible participants received compensation for their contribution. The study was approved by the Research Ethics Committees of Bournemouth University in the UK and Hamad Bin Khalifa University, Qatar.

Measures. The self-administered survey used in (Gurgun, Arden-Close, et al. 2023) was created in English. Two bilingual individuals translated it into Arabic using the back-translation process (Brislin, 1970). The questionnaire consisted of several primary sections, including an assessment of demographics (i.e., gender, age, and educational level). Moreover, the study examined the participant's likelihood of challenging the misinformation and how design techniques affected participants' willingness to counter misinformation by showing them eight high-fidelity prototypes grounded on the PSD in random order. The participants rated their likelihood of challenging the misinformation on a seven-point scale, ranging from "Extremely unlikely" to "Extremely likely," and how much each technique impacted them

to challenge misinformation on a seven-point scale, ranging from "Far too little" to "Far too much."

The assessment of personality traits was conducted using the 10-item Big Five Inventory (BFI-10) (Rammstedt and John, 2007), a widely accepted measure. The participants rated their agreement with each trait statement on a five-point Likert scale, ranging from "1: Strongly disagree" to "7: Strongly agree." A higher rating indicates a stronger presence of a trait. The BFI-10 provides reliable and valid measures of agreeableness, extraversion, conscientiousness, openness to experience, and neuroticism (Rammstedt and John, 2007). This scale includes two items for each trait, and the questionnaire items include questions such as "I see myself as someone who tends to find fault with others."

The study measured empathy and perspective-taking using the empathy and the perspective-taking subscales from the Interpersonal Reactivity Index (IRI) developed and validated by Davis (1980). The empathy subscale contains seven items to measure how much individuals feel warm, compassionate, and concerned for others. The perspective-taking also consists of seven items measuring how much individuals understand the psychological point of view of others. The questionnaire items include questions such as "I am often quite touched by things that I see happen" for empathy and "Before criticising somebody, I try to imagine how I would feel if I were in their place" for perspective-taking. The participants rated their agreement with each item on a five-point scale, ranging from "1: Does not describe me well" to "5: Describes me very well." The score for each subscale was obtained by averaging the responses to the respective items. In the current study, the internal reliability of the subscales was evaluated using Cronbach's Alpha. The empathy subscale demonstrated good reliability with an $\alpha = 0.82$ for the UK sample and an $\alpha = 0.61$, indicating acceptable reliability, for the Arab sample. Similarly, for the perspective-taking subscale, reliability was acceptable with an $\alpha = 0.79$ for the UK sample and $\alpha = 0.71$ for the Arab sample.

Data analysis. A combination of descriptive statistics, t-test, Mann-Whitney U test, Wilcoxon signed-rank test, and Binomial logistic regression were used to answer the research questions. Initially, descriptive statistics were performed to provide an overview of the demographic characteristics of the participants. A Welch's t-test was carried out to determine whether there was a difference in the willingness to challenge misinformation between the two cultures. It was used because the size and variance of the two samples were unequal. As the data of the perceived impact of the design techniques on willingness to challenge was not normally distributed, non-parametric tests were used. We examined the differences in the impact of the design interventions on users' willingness to challenge misinformation between the two cultures using the Mann-Whitney U test. Moreover, the Wilcoxon signed-rank test was used to investigate the differences in the impact of various design techniques on users' propensity to challenge misinformation compared to the standard comment box. Finally, binomial logistic regression was conducted to determine whether gender, age, empathy, perspective-taking, and personality traits might influence an individual's perception of the impact of design techniques, assessing them as being more or less persuasive than the standard comment box. The analysis was conducted using JASP software version 17.

Results
Descriptive Statistics
In this section, the demographic data of participants is summarised. The variables analysed included gender, age, and educational qualifications. In the UK sample, the demographic breakdown revealed that 104 females (41.6%), 143 males (57.2%), and three non-binary individuals (1.20%)

Table 2 Participant demographics.

		UK (*N* = 250)		Arab (*N* = 212)	
		Frequencies	%	Frequencies	%
Gender	Male	143	57.20	140	66.04
	Female	104	41.60	72	33.96
	Non-binary	3	1.20	0	0.00
Age	18–24	44	17.60	31	14.62
	25–34	94	37.60	113	53.31
	35–44	45	18.00	18	8.49
	Over 45	67	26.80	50	23.58
Education	University degree	157	62.80	169	79.72
	College degree	57	22.80	26	12.26
	Secondary and under	36	14.40	17	8.02

participated. In the Arabic sample, 72 females (33.96%) and 140 males (66.04%) participated. In terms of age, the UK sample consisted of 44 participants (17.6%) aged 18–24, 94 participants (37.6%) aged 25–34, 45 participants (18%) aged 35–44, and 67 participants (26.80%) over the age of 45. Similarly, the Arabic sample included 31 participants (14.62%) aged 18–24, 113 participants (53.31%) aged 25–34, 18 participants (8.49%) aged 35–44, and 50 participants (23.58%) over the age of 45. Regarding educational qualifications, most respondents in the UK sample (157 participants, 62.8%) held at least a university degree. This aligns with the Arabic sample, where 169 participants (79.72%) reported having a university degree. Additionally, in the UK sample, 57 participants (22.8%) had a college degree, which is comparable to the 26 participants (12.26%) with a college degree in the Arabic sample. Furthermore, 36 participants (14.4%) in the UK sample had completed secondary education, while the Arabic sample indicated 17 participants (8.02%) with secondary education or below. Detailed demographic information for each cultural group is shown in Table 2.

Willingness to challenge misinformation (UK vs. Arabs). The Welch's t-test was used to examine the differences in the willingness to challenge misinformation between the two cultural contexts, namely the UK and the Arabic contexts. The analysis of willingness to challenge misinformation was based on responses to the item of the questionnaire that assessed participants' likelihood of challenging misinformation. Participants were asked to recall a time when an acquaintance spread misinformation on Facebook. Then, they were asked how likely they would be willing to challenge misinformation publicly. The normality assumption check was not violated, and it was assessed using a Q-Q plot. The results revealed that the willingness to challenge misinformation was significantly higher in the Arab context (M = 4.70, SD = 1.80) than in the UK context (M = 3.30, SD = 1.94), with a significant t-value of 8.01 ($p < 0.001$). The effect size, as measured by Cohen's d, was 0.75, indicating a large-sized effect.

Perceived persuasiveness of design techniques (UK vs. Arabs). Given the participants' ratings on how much each of the eight high-fidelity prototypes of the design techniques (i.e., "predefined question stickers," "thinking face reaction," "private commenting," "sentence openers," "Fact checker badge," "social norm message," "tone detector," and "standard comment box") affects their willingness to challenge misinformation, the Mann-Whitney U test was used to examine the differences between the two cultural contexts in the effect of each of the design interventions

Table 3 Differences in perceived persuasiveness of design techniques between UK and Arab participants: Results from the Mann-Whitney U Test.

	UK (N = 250) M (SD)	Arab (N = 212) M (SD)	Mann-Whitney's U	p	r
Predefined question stickers	4.01 (1.81)	5.1 (1.95)	17338	<0.001	0.27
Thinking face reaction	4.49 (1.92)	4.92 (1.98)	22763	0.008	0.12
Private commenting	4.84 (1.76)	5.02 (2.09)	23757	0.051	0.09
Sentence openers	3.27 (1.67)	4.61 (1.91)	15757	<0.001	0.35
Fact checker badge	3.66 (1.87)	5.18 (1.98)	14954.5	<0.001	0.38
Social norm message	3.89 (1.68)	4.99 (1.93)	17132.5	<0.001	0.31
Tone detector	3.86 (1.79)	4.64 (2.08)	20162.5	<0.001	0.21
Standard comment box	3.66 (1.55)	4.33 (1.94)	20989.5	<0.001	0.18

on users' willingness to challenge misinformation. Table 3 presents the descriptive statistics and the results of the Mann-Whitney U test. The test results revealed significant cross-cultural differences in the perceived persuasiveness of design techniques between the UK and the Arab participants. The Arab participants rated the "predefined question stickers" (Mdn = 6), "sentence openers" (Mdn = 5), "fact checker badge" (Mdn = 6), "social norm message" (Mdn = 5), "tone detector" (Mdn = 5), and "standard comment box" (Mdn = 4) as significantly more persuasive than the UK participants, who rated them at Mdn = 4, 3, 4, 4, 4, and 4, respectively, all with $p < 0.001$. Rank-biserial (r) ranged from 0.18 to 0.38, indicating small to medium effects per Cohen's (1988) guidelines. The "thinking face reaction" was also found to be more persuasive among Arab participants (Mdn = 6) compared to UK participants (Mdn = 5) ($p = 0.008$), however, with a small effect size ($r = 0.12$). "Private commenting" was rated slightly higher among Arab participants (Mdn = 6) compared to UK participants (Mdn = 5) but did not attain statistical significance ($p = 0.051$), accompanied by a negligible effect size ($r = 0.09$), indicating a minimal difference in perceived persuasiveness between the two groups.

The Persuasiveness of Each Design Technique Versus the Standard Comment Box (SCB). Spearman's correlation revealed a minimal and nonsignificant association between the Standard Comment Box and the other techniques in the UK sample. The Arab sample showed more noticeable and significant associations between the "standard comment box" with "predefined question stickers," "sentence openers," "fact checker badge," "social norm message," and "tone detector" ($p < 0.001$ for all, except Tone Detector, where $p = 0.031$). However, these correlations were generally moderate. Table 4 provides a comprehensive overview of the correlation between all design techniques in both samples.

The Wilcoxon signed-rank test was used to examine the differences in the impact of each design technique on users' willingness to challenge misinformation compared to the standard comment box between the UK and the Arab cultural contexts, as shown in Table 5. In the UK sample, the "predefined question stickers," "thinking face reaction," and "private commenting" design techniques brought about statistically significant positive differences in their influence on willingness to challenge misinformation compared to the standard comment box with z-values of 2.43 ($p = 0.015$), 5 ($p < 0.001$), and 7.2 ($p < 0.001$), respectively. However, the "sentence openers" design technique yielded a significantly negative difference in its influence on willingness to challenge misinformation compared to the standard comment box ($z = -2.7$, $p = 0.007$).

In the Arabic sample, a similar trend was observed for "predefined question stickers," "thinking face reaction," and "private commenting," with z-values of -4.46 ($p < 0.001$), -3.37 ($p = 0.001$), and -3.64 ($p < 0.001$), respectively. Additionally, two further design techniques, "fact checker badge" and "social norm

message," also showed a significant positive difference in their influence on willingness to challenge misinformation compared to the standard comment box, with z-values of -4.92 ($p < 0.001$) and -4.12 ($p < 0.001$), respectively.

Excluding the UK's "sentence openers," both samples preferred PSD-informed design techniques over the standard comment box, indicating their superior influence on challenging misinformation.

Effects of Gender, Age, Empathy, Perspective-Taking, and Personality Traits. As "predefined question stickers," "private commenting," "thinking face reaction," "fact checker badge," "sentence openers," and "social norm messages" demonstrated statistically significant differences in their perceived persuasiveness on willingness to challenge within both or at least one of the cultural contexts compared to the standard comment box, we here investigated them further. Specifically, we aimed to determine whether gender, age, empathic concern of the IRI (i.e., empathy and perspective-taking), and BFI-10 personality traits (i.e., extroversion, conscientiousness, neuroticism, openness, and agreeableness) have an impact on these differences. We calculated the difference score for each technique by subtracting the standard comment box user rating from the user rating of the presented technique. We used a binary encoding system for the difference scores: '1' indicated a positive score, meaning that users rated the introduced design techniques as more persuasive than the standard comment box; '0' indicated a zero or negative score, meaning that users rated the introduced techniques as equally or less persuasive than the standard comment box.

We used binomial logistic regression models to examine the effects of gender, age, empathy, perspective-taking, and personality traits on the likelihood that participants rated the presented technique as more or less persuasive than the standard comment box. The encoded scores were the dependent variables. The linearity of the continuous variables with respect to the logit of the dependent variable was assessed via the Box-Tidwell (Box and Tidwell, 1962) procedure. All continuous independent variables were linearly related to the logit of the dependent variable. Hosmer and Lemeshow's tests indicated that the models for all design techniques fit the data well. Tables 6 and 7 show the binary logistic regression results for each design technique.

In the analysis of the significant techniques in both cultural contexts, several predictors of the persuasiveness of these design techniques were statistically significant. Concerning "predefined question stickers," extraversion (Odds Ratio = 1.25, $p = 0.027$) was significantly and positively associated with the impact of this technique in the Arabic sample, indicating that an increase in extraversion correlated with a higher likelihood of seeing "predefined question stickers" better than the "standard comment box" in persuading challenging misinformation. In the UK sample, age (Odds Ratio = 1.04, $p < 0.01$) and perspective-taking (Odds Ratio = 1.08, $p < 0.05$) were significantly positively

Table 4 Spearman's correlation between design techniques.

Design Techniques	1	2	3	4	5	6	7	8
1. Predefined question stickers	—	0.14* (0.047)	0.07 (0.309)	0.50*** (<0.001)	0.39*** (<0.001)	0.40*** (<0.001)	0.20** (0.003)	0.26*** (<0.001)
2. Thinking face reaction	0.08 (0.191)	—	0.21** (.002)	0.09 (0.168)	0.08 (0.259)	0.13 (0.053)	0.17* (0.013)	0.09 (0.175)
3. Private commenting	−0.03 (0.691)	−0.11 (0.073)	—	0.00 (0.965)	0.07 (0.288)	0.17* (0.015)	0.02 (0.731)	0.07 (0.345)
4. Sentence openers	0.34*** (<0.001)	0.09 (0.178)	0.01 (0.89)	—	0.34*** (<001)	0.42*** (<0.001)	0.31*** (<0.001)	0.29*** (<0.001)
5. Fact checker badge	0.34*** (<0.001)	0.02 (0.778)	0.05 (0.465)	0.28*** (<0.001)	—	0.43*** (<0.001)	0.20** (0.003)	0.27*** (<0.001)
6. Social norm message	0.25*** (<0.001)	0.12 (0.060)	0.00 (0.981)	0.30*** (<0.001)	0.31*** (<.001)	—	0.24** (0.001)	0.32*** (<0.001)
7. Tone detector	0.16** (0.009)	0.07 (0.288)	0.08 (0.211)	0.12 (0.055)	0.14* (.024)	0.21** (0.001)	—	0.15* (0.031)
8. Standard comment box	0.02 (0.710)	−0.06 (0.322)	0.11 (0.086)	0.08 (0.185)	0.07 (.257)	0.11 (0.094)	0.04 (0.479)	—

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.
The UK dataset is presented below the diagonal, and the Arab dataset is presented above the diagonal.

associated with "predefined question stickers," suggesting a similar pattern of increased influence of these factors on persuading challenging misinformation. Interestingly, a negative association was found between this technique and the openness to experience in the UK sample (Odds Ratio = 0.81, $p < 0.01$), unlike the Arabic sample, where this trait was not a significant predictor. Moreover, for the "thinking face reaction" technique, the analysis showed different significant predictors in each cultural context. In the Arabic sample, agreeableness was the only significant predictor (Odds Ratio = 0.79, $p = 0.019$), and it was negatively associated with the "thinking face reaction", indicating a decrease in the likelihood of the persuasive influence of this technique compared with the "standard comment box". In the UK sample, openness to experience was the sole significant predictor (Odds Ratio = 0.79, $p < 0.01$) and was also negatively associated with "thinking face reaction", suggesting a decrease in the persuasive influence by a factor of 0.21. Furthermore, it was shown that age was significantly negatively associated with "private commenting" in the UK sample only (Odds Ratio = 0.96, $p < 0.01$).

Regarding the analysis of the significant techniques in only one cultural context, for the "social norm messages," which was significant in the Arabic context only, openness to experience was significantly positively associated with it (Odds Ratio = 1.25, $p = 0.024$). In contrast, there were no significant predictors of the "sentence openers" and the "fact checker badge," although they had significant differences in persuading challenging misinformation compared to the standard comment box in UK and Arab contexts, respectively.

Discussion

As highlighted by Noman et al. (2024), certain factors, such as the perception of potential negative consequences (e.g., relationship cost, negative impact on the person being challenged, and futility of the correct action), along with injunctive norms, significantly influence individuals' willingness to challenge misinformation (Gurgun et al. 2024), with notable differences emerging between the UK and the Arab cultural contexts. The widespread nature of misinformation on social media platforms, combined with users' hesitancy to challenge it across diverse global cultures, necessitates a deeper exploration to motivate social media users to challenge it. Persuasive design techniques have been found to play a significant role in positively affecting users' willingness to challenge misinformation in the UK (Gurgun et al. 2023). Our study delved into the variances within the UK and Arab cultural contexts concerning the perception of the persuasiveness of the design interventions aimed at challenging misinformation proposed by Gurgun et al. (2023). The comparisons of the UK and the Arab samples revealed that except for the "private commenting," all other designs had significantly higher impact in the Arabic context compared to the UK context. This can be understood through the lens of high-context and low-context communication theories, such as Edward Hall's high-low context theory (Hall, 1976). According to Hall (1976), high-context cultures, like the Arabic culture, attribute significant importance to nonverbal communication and the situation's context. This preference aligns with the persuasive design techniques that we investigated, which largely rely on indirect communication forms. For example, the "thinking face reaction" and the "predefined question stickers" serve as nonverbal and contextually dependent cues for the users, which may explain their greater effectiveness in the Arab cultural context. Conversely, in the low-context UK culture, where explicit verbal communication is valued, these indirect design techniques might be less effective, as users tend to rely more on explicit message content rather than on nonverbal cues or on the context.

Table 5 The influence on willingness to challenge for each design intervention versus Standard Comment Box (SCB): Results from the Wilcoxon Signed-rank test.

Design Intervention	Ranks	UK (N = 250)					Arab (N = 212)				
		N	Mean Ranks	Sum of Ranks	Z	P	N	Mean Ranks	Sum of Ranks	Z	P
Predefined Stickers (PQS)	Neg.	80	110.84	8867.50	−2.430*	0.015	49	73.90	3621	−4.46*	<0.001
	Pos.	129	101.38	13077.50			107	80.61	8625		
	Ties	41					56				
Thinking Face Reaction (TFR)	Neg.	65	93.04	6047.50	−5.005*	<0.001	57	78.55	4477.5	−3.37*	0.001
	Pos.	136	104.81	14253.50			103	81.58	8402.5		
	Ties	49					52				
Private Commenting (PC)	Neg.	47	99.67	4684.50	−7.239*	<0.001	50	82.69	4134.50	−3.64*	<0.001
	Pos.	162	106.55	17260.50			107	77.28	8268.50		
	Ties	41					55				
Sentence Openers (SO)	Neg.	102	96.22	9814.00	−2.714¶	0.007	63	79.10	4983.50	−1.80*	0.073
	Pos.	76	80.49	6117.00			91	76.39	6951.50		
	Ties	72					58				
Fact Checker Badge (FCB)	Neg.	93	102.62	9543.50	−0.139*	0.890	41	64.07	2627.00	−4.92*	<0.001
	Pos.	103	94.78	9762.50			100	73.84	7384.00		
	Ties	54					71				
Social Norm Message (SNM)	Neg.	76	111.68	8488.00	−1.821*	0.060	50	64.42	3221.00	−4.12*	<0.001
	Pos.	123	92.78	11412.00			95	77.52	7364.00		
	Ties	51					67				
Tone Detector (TD)	Neg.	91	99.47	9051.50	−1.344*	0.170	61	82.46	5030.00	−1.82*	0.068
	Pos.	110	102.27	11249.50			94	75.11	7060.00		
	Ties	49					57				

Neg. = Negative Ranks (Design technique [e.g., PQS] impact < SCB impact); Pos. = Positive Ranks (Design technique [e.g., PQS] impact > SCB impact); * Based on negative ranks;
¶ Based on positive ranks.

The differences observed in the effectiveness of persuasive design techniques between the UK and Arab cultural contexts can also be explained by cultural attitudes towards confrontation. In Western cultures like the UK, which are characterised by high individualism, there is a greater value placed on open discussion and direct confrontation to resolve disputes or challenges, including confronting misinformation (Friedman et al. 2006). This cultural aspect supports the use of explicit and direct communication strategies, aligning with our findings that the investigated techniques rely on less confrontational forms, so they are less effective in the UK. Conversely, Arab culture, which tends to be more collectivistic, often adopts conflict avoidance strategies to preserve social harmony and relationships (Tjosvold and Sun, 2002). This preference for avoiding direct confrontation could explain why persuasive design techniques that utilise more subtle, indirect forms of communication, like those provided by "thinking face reactions" or "predefined question stickers," are more effective in this context. This inclination towards less confrontational design techniques aligns with high-context communication preferences, where much is left unsaid, and much is interpreted through contextual or non-verbal cues, reinforcing the need for design approaches that accommodate these cultural preferences in persuasion strategies.

Interestingly, the "private commenting" design technique was the exception to this general pattern, having an equivalent impact across both cultural contexts. This might be due to its inherent balance between indirect (i.e., private context) and direct (i.e., explicit commenting) communication. In this sense, it manages to cater to both high-context and low-context communication preferences, thereby overcoming cultural differences and universally encouraging the challenge of misinformation. This finding is in line with the concept of "convergence" in communication, where people can adjust their communication behaviours to decrease social differences (Dragojevic et al. 2015).

Building on the comparison between the persuasion of the design techniques in the UK and in an Arab cultural context, our exploration of their persuasiveness compared to the influence of the "standard comment box" underscored the differential influence of these design interventions in both cultures. The results of our study are generally consistent with those of the Gurgun et al. (2023) study, which was done in the UK context and found that the "predefined question stickers," "thinking face reaction," and "private commenting" influenced the users positively to challenge misinformation. These techniques were rated as more persuasive than the standard comment box, indicating that people preferred the system to provide them with ready-made options to facilitate challenging misinformation rather than having to compose comments themselves.

However, there were some differences between the two cultural contexts. In the UK, the "sentence openers" was less persuasive than the standard comment box, whereas it had no significant impact on Arab users. Previous research has shown that users are more likely to accept and use information technologies that they perceive as easy to use (Venkatesh, 2000). However, based on our findings, simply providing guidance without cultural consideration may not motivate users to challenge misinformation. Even though participants were offered the "sentence openers," developing their arguments might require more effort, particularly in low-context cultures such as the UK, where communication is explicit and largely dependent on words (Hall, 1976). Conversely, the strategy of making the misinformation challenge easier by providing a text to be completed might not have a significant effect in high-context cultures such as the Arab culture, which tend to rely heavily on nonverbal cues such as body language, tone, and the overall context (Hall, 1976). This finding echoes research that suggests the importance of considering cultural factors for effective technology design and usage (Marcus and Gould, 2000).

Table 6 Binomial Logistic Regressions predicting the difference score between the persuasiveness of Standard Comment Box and the persuasiveness of the presented design techniques (higher/lower).

		PQS		TFR		PC	
		UK	Arab	UK	Arab	UK	Arab
Age	B	0.04*	0.04	−0.01	0.00	−0.04*	0.00
	SE	0.01	0.02	0.01	0.02	0.01	0.02
	Odds Ratio	1.04	1.04	0.99	1.00	0.96	1.00
Gender (Female)	B	−0.38	−0.29	−0.45	0.01	−0.15	0.09
	SE	0.31	0.31	0.30	0.31	0.31	0.31
	Odds Ratio	0.68	0.75	0.64	1.01	0.86	1.09
Extraversion	B	0.11	0.22*	0.01	0.00	0.06	0.04
	SE	0.08	0.10	0.08	0.10	0.08	0.09
	Odds Ratio	1.12	1.25	1.01	1.00	1.06	1.04
Agreeableness	B	−0.18	−0.09	0.01	−0.23*	0.06	−0.13
	SE	0.10	0.10	0.09	0.10	0.10	0.10
	Odds Ratio	0.84	0.91	1.01	0.79	1.07	0.88
Conscientiousness	B	−0.12	−0.06	0.06	0.00	−0.09	0.01
	SE	0.10	0.10	0.10	0.10	0.10	0.10
	Odds Ratio	0.89	0.94	1.07	1.00	0.91	1.01
Neuroticism	B	−0.02	0.16	0.03	0.07	0.01	0.12
	SE	0.08	0.08	0.08	0.08	0.08	0.08
	Odds Ratio	0.98	1.17	1.03	1.07	1.01	1.13
Openness to experience	B	−0.21*	−0.01	−0.24*	0.02	−0.09	−0.05
	SE	0.08	0.10	0.08	0.10	0.08	0.10
	Odds Ratio	0.81	0.99	0.79	1.02	0.91	0.95
Empathy	B	−0.03	−0.01	0.01	−0.02	0.04	0.03
	SE	0.04	0.04	0.04	0.04	0.04	0.04
	Odds Ratio	0.97	0.99	1.01	0.98	1.04	1.03
Perspective Taking	B	0.08**	0.03	−0.01	0.00	0.00	−0.01
	SE	0.04	0.04	0.04	0.04	0.04	0.04
	Odds Ratio	1.08	1.03	0.99	1.00	1.00	0.99
Constant	B	1.14	−2.50	1.75	1.34	2.06	−0.05
	SE	1.54	1.58	1.53	1.55	1.60	1.54
	Odds Ratio	3.14	0.08	5.75	3.82	7.83	0.95
Modal Summary	Hosmer and Lemeshow X^2	5.28	7.02	9.87	4.23	9.22	7.71
	df	8	8	8	8	8	8
	p	0.726	0.534	0.270	0.836	0.320	0.463
	Nagelkerke R^2	0.11	0.07	0.07	0.06	0.10	0.04

B unstandardised regression coefficient, *SE* standard error, *df* Degrees of Freedom, *p* significance level, *p < 0.05, **p < 0.01.

On the other hand, in the Arabic context only, the "fact checker badge" and the "social norm message" were more persuasive than the standard comment box. The "fact checker badge" seems to operate at the intersection of recognition, social conformity, and cultural communication preferences, thereby providing a persuasive tool specifically attuned to high-context cultures, such as the Arabic culture. High-context cultures often rely heavily on nonverbal cues and on the context of communication (Hall, 1976). In such settings, the "fact checker badge" may serve as a symbolic recognition for users who challenge misinformation, thus tapping into the inherent desire for social approval (Leary and Kowalski, 1990). This outcome highlights the efficiency of badges in high-context cultures, where communication is deeply intertwined with social relationships and shared understanding. They do not only serve as a signal of individual achievement but also underline a shared commitment within the community towards challenging misinformation. Similarly, the "social norm message" appears to cater effectively to users in high-context cultures because of their heightened sensitivity to social cues and group norms. For instance, normative messages informing social media users about the prevalent acceptance of challenging misinformation in the community can create a sense of social pressure and expectations to conform to this majority behaviour (Andı and Akesson, 2020). This aligns with research studies that have shown the role of social norms in impacting a

person's behaviour (Melnyk et al. 2022). The normative message may also implicitly suggest that challenging misinformation is valuable and rewarding, with potential benefits in enhancing one's social standing and reputation within the community. Consequently, leveraging the potent force of social influence, the "social norm message" can bolster users' motivation and confidence in challenging misinformation.

We aimed to discern the potential impact of individual characteristics such as gender, age, empathy, perspective-taking, and personality traits on the influence of design techniques that persuaded people more than the standard comment box across the UK and Arab cultural contexts. The findings from the binomial logistic regression models revealed different significant predictors for each technique, thereby providing valuable insights into the multifaceted nature of the persuasiveness of design techniques. Considering the "predefined question stickers," different influential factors emerged in the UK and in the Arab cultures. In the UK, age and perspective-taking serve as positive influencers, suggesting that older individuals and those who frequently engage in perspective-taking are more likely to respond positively to this design technique. However, openness, which is generally associated with active imagination and experience-seeking (Rammstedt and John, 2007), appears to be a negative influencer in the UK, indicating that these individuals may prefer more personalised or less structured ways of

Table 7 Binomial Logistic Regressions predicting the difference score between the persuasiveness of Standard Comment Box and the persuasiveness of the presented design techniques (higher/lower).

		SO (UK only)	FCB (Arab only)	SNM (Arab only)
Age	B	0.00	0.01	0.02
	SE	0.01	0.02	0.02
	Odds Ratio	1.00	1.01	1.02
Gender (Female)	B	−0.50	−0.07	−0.03
	SE	0.32	0.30	0.31
	Odds Ratio	0.61	0.93	0.97
Extraversion	B	−0.06	−0.05	−0.11
	SE	0.09	0.09	0.10
	Odds Ratio	0.94	0.96	0.89
Agreeableness	B	0.02	−0.14	−0.16
	SE	0.10	0.10	0.10
	Odds Ratio	1.02	0.87	0.85
Conscientiousness	B	−0.13	−0.03	−0.07
	SE	0.10	0.10	0.10
	Odds Ratio	0.88	0.97	0.93
Neuroticism	B	−0.04	0.02	0.05
	SE	0.08	0.08	0.08
	Odds Ratio	0.96	1.02	1.05
Openness to experience	B	−0.14	0.09	0.22*
	SE	0.08	0.09	0.10
	Odds Ratio	0.87	1.10	1.25
Empathy	B	0.02	0.03	0.00
	SE	0.05	0.04	0.04
	Odds Ratio	1.02	1.03	1.00
Perspective Taking	B	0.04	0.00	0.04
	SE	0.04	0.04	0.04
	Odds Ratio	1.04	1.00	1.04
Constant	B	0.66	−0.43	−1.34
	SE	1.62	1.53	1.57
	Odds Ratio	1.93	0.65	0.26
Modal Summary	Hosmer and Lemeshow X^2	4.83	9.41	4.96
	Df	8	8	8
	p Value	0.770	0.309	0.761
	Nagelkerke R^2	0.05	0.03	0.07

B unstandardised regression coefficient, *SE* standard error, *df* Degrees of Freedom, *p* significance level, *∗p* < 0.05, *∗∗p* < 0.01.
The table contains techniques that had a significant impact on users' willingness to challenge misinformation in one culture only.

challenging misinformation. In the Arab context, extraversion emerges as a positive influencer, which could be indicative of extraverted individuals' propensity for social engagement and willingness to use this technique as a structured communication tool. The "thinking face reaction" design technique also yielded varying responses across the two cultures. In the UK, individuals high in openness were less likely to be positively influenced by this technique, perhaps because people who are more open to experience might prefer a more nuanced form of expressing their thoughts or doubts. However, in the Arab context, agreeableness has surfaced as a negative influencer. This finding suggests that more agreeable individuals, often characterised by their tendency for harmony and dislike of confrontations (Jensen-Campbell and Graziano, 2001), might find it challenging to use a design feature that signifies doubt or disagreement.

Moreover, considering the design intervention of "private commenting," age emerged as an influential factor in the UK context only. This design technique was more persuasive for younger adults, who preferred privately confronting misinformation. This finding reflects that young people are more concerned than older people about how they appear to others and are less willing to express their views publicly (Henry et al. 2018). In contrast to our expectations, the "sentence openers" and the "fact checker badge" did not demonstrate any significant influencers in either culture, indicating the need for more in-depth research to understand their potential applicability better. Lastly, in the "social norm message" case, openness was observed as a positive influencer within the Arab context only. A possible explanation for this finding may be drawn from a previous study that associated individuals' openness with understanding and accepting of diverse viewpoints (John and Srivastava, 1999). As such, people who are high in openness might be more likely to accept and internalise normative messages, even if they diverge from their current beliefs or practices.

The results from our study present several notable implications for research, policy, and practice in the fields of misinformation, social media design, and cross-cultural communication. We argue that our research contributes to the existing body of knowledge in these fields by offering a nuanced understanding of how cultural communication preferences impact the effectiveness of persuasive design techniques. Furthermore, our study reinforces the value of considering high-context and low-context communication theories, such as Hall's high-low context theory (Hall, 1976), in cross-cultural research. Our findings also highlight the influence of individual characteristics, such as gender, age, empathy, perspective-taking, and personality traits, on the effectiveness of different persuasive design techniques, thereby adding a layer of complexity to communication and persuasion theories. This approach is consistent with the previous studies in cross-cultural human-computer interaction (HCI), which highlight the importance of cultural factors for designing effective user-centric interfaces (Adnan et al. 2020; Reinecke and Bernstein, 2011).

Moreover, the findings of this study have practical implications for social media platform developers and designers who aim to combat misinformation. By considering cultural differences in communication, it is possible to customise and adapt these persuasive techniques to the communication styles of the targeted users and optimise their effectiveness. For instance, our study suggests that nonverbal cues, like the "thinking face reaction," might be more effective in high-context cultures such as the Arabic culture, while techniques like "private commenting" might work equally well across both high-context and low-context cultures. However, total privacy on social media is undesirable since it could lead to unrestrained behaviour (Lapidot-Lefler and Barak, 2012). Thus, a balance can be established by implementing semi-private solutions that protect the user and the person posting the misinformation. Furthermore, our results support the idea that policymakers can leverage insights from our study to inform regulations concerning the challenge of misinformation on social media platforms. As a practical application, guidelines could be developed that encourage social media platforms to embed cultural preferences in their design approaches aimed at combating misinformation, which may not only enhance the effectiveness of the interventions but also nurture user trust. This could be a step forward for enhancing the accuracy and credibility of the information shared on social media platforms (Vijaykumar et al. 2021). This aligns with the findings of Wang et al. (2016) and (Wang et al. 2022), reinforcing the importance of trust in shaping user behaviour and interaction with online platforms, which can be a promising approach to curb the spread of misinformation.

Limitations and Future Research

This study primarily focused on social media users who are actively engaging in public discourse, which may not reflect the perspectives of passive users or those who sparingly use social media, potentially limiting the generalizability of our findings. Furthermore, we concentrated on addressing misinformation among Facebook acquaintances. Nevertheless, the notion might encompass various degrees of relational closeness. Future studies should investigate the influence of relational closeness and cultural context on individuals' disinclination to challenge misinformation. Regarding methodological limitations, our study relied on self-reported measures of the perceived effectiveness of various design interventions. Although useful for gathering subjective perceptions, self-reported data can sometimes suffer from biases, such as social desirability or memory recall (Podsakoff et al. 2003). Therefore, the self-reported nature of our data may have affected the accuracy of participants' responses.

Moreover, while our analysis was fruitful in revealing cultural differences between the UK and Arab cultural contexts, it overlooked within-culture variations that might influence these outcomes. For instance, within each cultural context, differences based on regional, socioeconomic, or other demographic factors (Hepp, 2015) might exist, which our study did not address. In addition, we specifically focused on challenging misinformation within the context of social media. While this is a significant and timely issue, it also limits our understanding of how these persuasive design techniques might be effective (or not) in other communication contexts. Finally, we examined participant responses at a single point in time. Given the dynamic nature of social media and societal norms, longitudinal studies could provide more insights into how changes over time might affect the strategies used to combat misinformation.

Future research can build upon our findings, addressing the limitations mentioned, and advancing the understanding of challenging misinformation. For instance, while we attempted to gauge the persuasiveness of various design techniques through participants' self-reported willingness to challenge misinformation, this indirect approach does not directly measure actual behaviour. Echoing the concerns of behavioural researchers (Berman, 2018), this recommendation points towards the need for future research to implement more direct behavioural measures by employing different settings, such as experimental or observational ones, for a more accurate evaluation of the effectiveness of these design interventions. Moreover, future research should delve deeper into the individual characteristics that influence the effectiveness of persuasive design interventions.

In addition, future research should explore the development and evaluation of innovative persuasive design techniques that are specifically tailored to align with high and low-context cultural communications. This endeavour will expand the understanding of cultural communication styles on digital platforms and contribute to more effective and inclusive strategies for mitigating the spread of misinformation globally. This supports research that suggests developing context-sensitive user interfaces (Shen et al. 2006; Zimmermann et al. 2014) and tailoring communication means to the target group's national culture (Marcus and Gould, 2000). To further enhance the impact of persuasive design techniques tailored for high and low-context cultural communications, it is recommended to integrate gamification elements, such as points and badges, into the design. This enhancement aligns with the principles of increasing user engagement and interaction, as outlined by research in gamification (Deterding et al. 2011), and could be crucial in designing more innovative design techniques for combating misinformation across diverse cultural landscapes. Design architectural changes known as "Nudges" have also been shown to influence user behaviour (Caraban et al. 2019). It is crucial to explore using nudges to build innovative design techniques that facilitate the engagement of people from different cultural backgrounds in challenging misinformation. For example, in high-context cultures, incorporating visual cues, symbols, and context-specific references into the design techniques can better resonate with users' communication patterns and nonverbal communication preferences. On the other hand, in low-context cultures, design techniques may prioritise clear and concise messaging to appeal to individuals and foster their active participation in combating misinformation. In line with studies demonstrating the role of customer co-creation behaviour in developing innovative services (Moghadamzadeh et al. 2020), future research might investigate how social media users' co-creation behaviour contributes to the development of innovative persuasive design techniques in various cultural contexts.

Further research should explore how our findings apply to diverse cultural contexts beyond the UK and Arab cultures. This suggestion resonates with the literature's emphasis on the value of cross-cultural research on communication through social media platforms (Yuna et al. 2022). Finally, when social media users challenge other people's behaviour online, including posting misinformation, it may put their personal relationships at risk. People's perceptions of risk can vary, with some leaning towards risk-taking and others towards risk-avoidance. This inclination might seem to be influenced by cognitive ability (Amador-Hidalgo et al. 2021). Individuals with higher abilities might be more inclined to challenge, while those with lower abilities might refrain. Future studies should investigate the effects of the need for cognition on the propensity to challenge misinformation.

Conclusion

Our study aimed to test the effectiveness of various persuasive design techniques in encouraging social media users to challenge misinformation in UK and Arab cultural contexts and to identify the individual characteristics that predict this influence. The design techniques used in this study are predefined question stickers, thinking face reaction, sentence openers, fact checker badge, social norm messages, tone detector, and private commenting. The techniques exemplify seven persuasive principles: suggestion, liking, tunnelling, recognition, normative influence, self-monitoring, and reduction, respectively. The findings demonstrate that, except for the private commenting, all techniques had a significantly greater impact in the Arab context compared to the UK. Further analysis showed that techniques like thinking face reactions were more effective than the standard comment box in both cultures, while the sentence openers technique was less effective only in the UK context, and some techniques, such as the fact checker badge, had a greater impact only in the Arab context. The core findings of our study indicate the role of cultural communication preferences, as dictated by high-context and low-context cultures, in influencing the effectiveness of these techniques. Additionally, individual characteristics, such as age and personality traits, also significantly impact the persuasiveness of these design techniques. Our work is a pioneering effort to understand the effectiveness of persuasive design techniques in challenging misinformation within the contrasting contexts of the UK and Arab cultures. We have brought the often-overlooked non-Western perspective, specifically the Arab cultural context, to the forefront of this global issue, thereby contributing significantly to the literature. Thus, our research underscores the need for social media platforms to adopt a culturally aware design approach to effectively address

misinformation. Through a deeper understanding and incorporation of these cultural and individual nuances, we can pave the way for developing more resonant and practical strategies to curb the global issue of misinformation on social media.

Data availability

References

Adnan, M, Wei, WC, & Ghazali, M (2020). Localization and Globalization of Website Design: A pilot study focuses on comparison of government websites. *Proceedings of the 2020 Symposium on Emerging Research from Asia and on Asian Contexts and Cultures*, 2–2. https://doi.org/10.1145/3391203.3391212

Aïmeur E, Amri S, Brassard G (2023) Fake news, disinformation and misinformation in social media: a review. Social Network Analysis and Mining 2023 13:1 13(1):1–36

Alqahtani F, Meier S, Orji R (2022) Personality-based approach for tailoring persuasive mental health applications. Use. Modeling Use.-Adapted Interact. 32(3):253–295. https://doi.org/10.1007/S11257-021-09289-5

Amador-Hidalgo L, Brañas-Garza P, Espín AM, García-Muñoz T, Hernández-Román A (2021) Cognitive abilities and risk-taking: errors, not preferences. Eur. Economic Rev. 134:103694. https://doi.org/10.1016/J.EUROECOREV.2021.103694

Andı S, Akesson J (2020) Nudging away false news: evidence from a social norms experiment. Digital Journalism 9(1):106–125. https://doi.org/10.1080/21670811.2020.1847674

Bao T, Wei L, Yu Y (2022) The impact of information interventions on public opinion on social media regulation: evidence from a survey on Twitter's Trump Ban. J. Behav. Exp. Econ. 101:101947. https://doi.org/10.1016/J.SOCEC.2022.101947

Berman GJ (2018) Measuring behavior across scales. BMC Biology16:1 16(1):1–11

Blau I, Barak A (2012) How do personality, synchronous media, and discussion topic affect participation? J. Educ. Technol. Soc. 15(2):12–24

Bode L, Vraga EK (2015) In Related News, That was Wrong: The Correction of Misinformation Through Related Stories Functionality in Social Media. J Commun. 65(4):619–638. https://doi.org/10.1111/JCOM.12166

Bode L, Vraga EK (2017) See Something, Say Something: Correction of Global Health Misinformation on Social Media. Health Commun. 33(9):1131–1140. https://doi.org/10.1080/10410236.2017.1331312

Box GEP, Tidwell PW (1962) Transformation of the Independent Variables. Technometrics 4(4):531–550. https://doi.org/10.1080/00401706.1962.10490038

Brislin RW (1970) Back-translation for cross-cultural research. J. Cross-Cultural Psychol. 1(3):185–216

Caraban, A, Técnico, IS, Karapanos, E, Gonçalves, D, & Campos, P (2019). 23 ways to nudge: A review of technology-mediated nudging in human-computer interaction. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–15. https://doi.org/10.1145/3290605.3300733

Cemiloglu D, Arden-Close E, Hodge SE, Ali R (2023) Explainable persuasion for interactive design: The case of online gambling. J. Syst. Softw. 195:111517. https://doi.org/10.1016/J.JSS.2022.111517

Chadwick, A, Vaccari, C, & Hall, NA (2023). What Explains the Spread of Misinformation in Online Personal Messaging Networks? Exploring the Role of Conflict Avoidance. *Digital Journalism*, 1–20. https://doi.org/10.1080/21670811.2023.2206038

Choraś M, Demestichas K, Giełczyk A, Herrero Á, Ksieniewicz P, Remoundou K, Urda D, Woźniak M (2021) Advanced Machine Learning techniques for fake news (online disinformation) detection: A systematic mapping study. Appl. Soft Comput. 101:107050. https://doi.org/10.1016/J.ASOC.2020.107050

Cohen, J (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum

Cook J, Lewandowsky S, Ecker UKH (2017) Neutralizing misinformation through inoculation: Exposing misleading argumentation techniques reduces their influence. PLOS ONE 12(5):e0175799. https://doi.org/10.1371/JOURNAL.PONE.0175799

Davis, MH (1980). A multidimensional approach to individual differences in empathy. *Journal of Personality and Social Psychology*, 10(85)

Deterding, S, Dixon, D, Khaled, R, & Nacke, L (2011). From game design elements to gamefulness: Defining "gamification". *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments*, 9–15. https://doi.org/10.1145/2181037.2181040

Dragojevic, M, Gasiorek, J, & Giles, H (2015). Communication Accommodation Theory. *The International Encyclopedia of Interpersonal Communication*, 1–21. https://doi.org/10.1002/9781118540190.WBEIC006

Elaheebocus SMRA, Weal M, Morrison L, Yardley L (2018) Peer-based social media features in behavior change interventions: systematic review. J. Med. Internet Res. 20(2):e8342. https://doi.org/10.2196/JMIR.8342

Eysenck, M, & Keane, M (2015). *Cognitive psychology: A student's handbook* (7th ed.). Psychology Press. https://doi.org/10.4324/9781315778006

Ford, G, & Kotzé, P (2005). Designing usable interfaces with cultural dimensions. In MF Costabile & F Paterno (Eds.), *Human-Computer Interaction - INTERACT 2005: Vol. 3585 LNCS* (pp. 713–726). Springer, Berlin, Heidelberg. https://doi.org/10.1007/11555261_57

Friedman R, Chi SC, Liu LA (2006) An expectancy model of Chinese-American differences in conflict-avoiding. J. Int. Bus. Stud. 37(1):76–91. https://doi.org/10.1057/PALGRAVE.JIBS.8400172

Gurgun, S, Arden-Close, E, McAlaney, J, Phalp, K, & Ali, R (2023). Can We Redesign Social Media to Persuade People to Challenge Misinformation? An Exploratory Study. *Persuasive Technology: 18th International Conference, PERSUASIVE 2023, Eindhoven, The Netherlands, April 19--21, 2023, Proceedings*, 123–141. https://doi.org/10.1007/978-3-031-30933-5_9

Gurgun, S, Arden-Close, E, Phalp, K, & Ali, R (2022). Online silence: why do people not challenge others when posting misinformation? *Internet Research*, ahead-of-print(ahead-of-print). https://doi.org/10.1108/INTR-06-2022-0407

Gurgun S, Cemiloglu D, Close EA, Phalp K, Nakov P, Ali R (2024) Why do we not stand up to misinformation? Factors influencing the likelihood of challenging misinformation on social media and the role of demographics. Technol. Soc. 76:102444. https://doi.org/10.1016/J.TECHSOC.2023.102444

Halko, S, & Kientz, JA (2010). Personality and persuasive technology: An exploratory study on health-promoting mobile applications. *Persuasive Technology: 5th International Conference, 6137 LNCS*, 150–161. https://doi.org/10.1007/978-3-642-13226-1_16

Hall, E (1976). *Beyond culture*. Anchor

Henrich J, Heine SJ, Norenzayan A (2010) Most people are not WEIRD. Nature 466(7302):29–29. https://doi.org/10.1038/466029a

Henry JD, von Hippel W, Nangle MR, Waters M (2018) Age and the experience of strong self-conscious emotion. Aging Ment. Health 22(4):497–502. https://doi.org/10.1080/13607863.2016.1268094

Hepp, A (2015). *Transcultural communication*. John Wiley & Sons

Hofstede, G (2001). *Culture's consequences: Comparing values, behaviors, institutions and organizations across nations* (2nd ed.). Sage publications

Hummel D, Maedche A (2019) How effective is nudging? A quantitative review on the effect sizes and limits of empirical nudging studies. J. Behav. Exp. Econ. 80:47–58. https://doi.org/10.1016/J.SOCEC.2019.03.005

Jensen-Campbell LA, Graziano WG (2001) Agreeableness as a moderator of interpersonal conflict. J. Personal. 69(2):323–362. https://doi.org/10.1111/1467-6494.00148

John O, Srivastava S (1999) The Big-Five trait taxonomy: History, measurement, and theoretical perspectives. Handb. Personality: Theory Res. 2:102–138

Jost, PJ, Pünder, J, & Schulze-Lohoff, I (2020). Fake news - Does perception matter more than the truth? *Journal of Behavioral and Experimental Economics*, 85. https://doi.org/10.1016/J.SOCEC.2020.101513

Kligler-Vilenchik N (2022) Collective social correction: addressing misinformation through group practices of information verification on WhatsApp. Digital Journalism 10(2):300–318. https://doi.org/10.1080/21670811.2021.1972020

Kuklinski JH, Quirk PJ, Jerit J, Schwieder D, Rich RF (2000) Misinformation and the Currency of Democratic Citizenship. J. Politics 62(3):790–816. https://doi.org/10.1111/0022-3816.00033

Lapidot-Lefler N, Barak A (2012) Effects of anonymity, invisibility, and lack of eye-contact on toxic online disinhibition. Computers Hum. Behav. 28(2):434–443. https://doi.org/10.1016/J.CHB.2011.10.014

Leary MR, Kowalski RM (1990) Impression Management: A Literature Review and Two-Component Model. Psychological Bull. 107(1):34–47. https://doi.org/10.1037/0033-2909.107.1.34

Lewandowsky S, Ecker UKH, Seifert CM, Schwarz N, Cook J (2012) Misinformation and Its Correction: Continued Influence and Successful Debiasing. Psychological Sci. Public Interest, Suppl. 13(3):106–131. https://doi.org/10.1177/1529100612451018

Marcus A, Gould EW (2000) Crosscurrents: cultural dimensions and global Web user-interface design. Interactions 7(4):32–46. https://doi.org/10.1145/345190.345238

Matyszczyk, C (2014, May 27). "Asteroid to hit Earth" story goes up on CNN, people believe it. https://www.cnet.com/science/cnn-posts-asteroid-to-hit-earth-article-people-take-it-seriously/

McGowan A, Sittig S, Bourrie D, Benton R, Iyengar S (2022) The Intersection of Persuasive System Design and Personalization in Mobile Health: Statistical Evaluation. JMIR MHealth UHealth 10(9):e40576. https://doi.org/10.2196/40576

Melnyk V, Carrillat FA, Melnyk V (2022) The Influence of Social Norms on Consumer Behavior: A Meta-Analysis. J. Mark. 86(3):98–120. https://doi.org/10.1177/00222429211029199

Moghadamzadeh A, Ebrahimi P, Radford S, Salamzadeh A, Khajeheian D (2020) Investigating the role of customer co-creation behavior on social media platforms in rendering innovative services. Sustainability 12, 6926

Morgan, JC, & Cappella, JN (2023). The Effect of Repetition on the Perceived Truth of Tobacco-Related Health Misinformation Among U.S. Adults. *Journal of Health Communication*, 28(3). https://doi.org/10.1080/10810730.2023.2192013

Noman, M, Gurgun, S, Phalp, K, Nakov, P, & Ali, R (2024). Challenging Others When Posting Misinformation: A UK vs. Arab Cross-Cultural Comparison on the Perception of Negative Consequences and Injunctive Norms. *Behaviour and Information Technology*, 1–21. https://doi.org/10.1080/0144929X.2023.2298306

Nyhan B, Reifler J (2010) When corrections fail: The persistence of political misperceptions. Political Behav. 32(2):303–330. https://doi.org/10.1007/S11109-010-9112-2

Oinas-Kukkonen H (2013) A foundation for the study of behavior change support systems. Personal. Ubiquitous Comput. 17(6):1223–1235. https://doi.org/10.1007/S00779-012-0591-5

Oinas-Kukkonen H, Harjumaa M (2009) Persuasive Systems Design: Key Issues, Process Model, and System Features. Communications of the Association for Information Systems 24(1)):28

Pan, W, Liu, D, & Fang, J (2021) An Examination of Factors Contributing to the Acceptance of Online Health Misinformation. Frontiers in Psychology, 12. https://doi.org/10.3389/FPSYG.2021.630268

Podsakoff PM, MacKenzie SB, Lee J-Y, Podsakoff NP (2003) Common method biases in behavioral research: a critical review of the literature and recommended remedies. J. Appl. Psychol. 88(5):879–903

Rammstedt B, John OP (2007) Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German. J. Res. Personal. 41:203–212. https://doi.org/10.1016/j.jrp.2006.02.001

Reinecke K, Bernstein A (2011) Improving performance, perceived usability, and aesthetics with culturally adaptive user interfaces. ACM Trans. Computer-Hum. Interact. (TOCHI) 18(2):29. https://doi.org/10.1145/1970378.1970382

Shen ST, Woolley M, Prior S (2006) Towards culture-centred design. Interact. Computers 18(4):820–852. https://doi.org/10.1016/J.INTCOM.2005.11.014

Tandoc Jr EC, Lim D, Ling R (2020) Diffusion of disinformation: How social media users respond to fake news and why. Journalism 21(3):381–398. https://doi.org/10.1177/1464884919868325

Tjosvold D, Sun HF (2002) Understanding conflict avoidance: Relationship, motivations, actions, and consequences. Int. J. Confl. Manag. 13(2):142–164. https://doi.org/10.1108/eb022872

Tversky A, Kahneman D (1974) Judgment under uncertainty: Heuristics and biases. Science 185(4157):1124–1131. https://doi.org/10.1126/SCIENCE.185.4157.1124

Valenzuela S, Kim Y, Gil De Zúñiga H (2012) Social Networks that Matter: Exploring the Role of Political Discussion for Online Political Participation. Int. J. Public Opin. Res. 24(2):163–184. https://doi.org/10.1093/ijpor/edr037

Venkatesh V (2000) Determinants of Perceived Ease of Use: Integrating Control, Intrinsic Motivation, and Emotion into the Technology Acceptance Model. Inf. Syst. Res. 11(4):342–365. https://doi.org/10.1287/ISRE.11.4.342.11872

Vicari R, Komendatova N (2023) Systematic meta-analysis of research on AI tools to deal with misinformation on social media during natural and anthropogenic hazards and disasters. Humanities Soc. Sci. Commun. 10(1):332. https://doi.org/10.1057/s41599-023-01838-0

Vicario MDel, Bessi A, Zollo F, Petroni F, Scala A, Caldarelli G, Stanley HE, Quattrociocchi W (2016) The spreading of misinformation online. Proc. Natl Acad. Sci. USA 113(3):554–559. https://doi.org/10.1073/pnas.1517441113

Vijaykumar S, Jin Y, Rogerson D, Lu X, Sharma S, Maughan A, Fadel B, de Oliveira Costa MS, Pagliari C, Morris D (2021) How shades of truth and age affect responses to COVID-19 (Mis)information: randomized survey experiment among WhatsApp users in UK and Brazil. Humanities Soc. Sci. Commun. 8(1):88. https://doi.org/10.1057/s41599-021-00752-7

Vosoughi S, Roy D, Aral S (2018) The spread of true and false news online. Science 359(6380):1146–1151. https://doi.org/10.1126/SCIENCE.AAP9559

Walter N, Murphy ST (2018) How to unring the bell: A meta-analytic approach to correction of misinformation. Commun. Monogr. 85(3):423–441. https://doi.org/10.1080/03637751.2018.1467564

Wang, J, Shahzad, F, Ahmad, Z, Abdullah, M, & Hassan, NM (2022). Trust and consumers' purchase intention in a social commerce platform: A meta-analytic approach. *Sage Open*, 12(2). https://doi.org/10.1177/21582440221091262

Wang Y, Min Q, Han S (2016) Understanding the effects of trust and risk on individual behavior toward social media platforms: A meta-analysis of the empirical evidence. Computers Hum. Behav. 56:34–44. https://doi.org/10.1016/J.CHB.2015.11.011

Weinmann M, Schneider C, vom Brocke J (2016) Digital Nudging. Bus. Inf. Syst. Eng. 58(6):433–436. https://doi.org/10.1007/S12599-016-0453-1

Wiafe I, Koranteng FN, Owusu E, Ekpezu AO, Gyamfi SA (2020) Persuasive social features that promote knowledge sharing among tertiary students on social networking sites: An empirical study. J. Computer Assist. Learn. 36(5):636–645. https://doi.org/10.1111/JCAL.12433

Yuna D, Xiaokun L, Jianing L, Lu H (2022) Cross-Cultural Communication on Social Media: Review From the Perspective of Cultural Psychology and Neuroscience. Front. Psychol. 13:858900. https://doi.org/10.3389/FPSYG.2022.858900

Zimmermann, G, Vanderheiden, GC, & Strobbe, C (2014). Towards deep adaptivity - A framework for the development of fully context-sensitive user interfaces. In C Stephanidis & M Antona (Eds.) Universal Access in Human-Computer Interaction. Design and Development Methods for Universal Access: Vol. 8513 LNCS (Issue PART 1, pp. 299–310). Springer. https://doi.org/10.1007/978-3-319-07437-5_29

Acknowledgements

Author contributions

MN was responsible for preparing the study for the Arab context, conducting data collection and analysis, interpreting the findings, and writing the paper. SG conceptualised the research, conducted the original study in the UK, assisted with data preparation and analysis, and contributed to the interpretation of the results. KP provided valuable feedback, reviewed the paper, and made revisions. RA was involved in all stages of the research and provided supervision. All authors have reviewed and approved the final version of the manuscript.

Competing interests

The authors declare no competing interest.

Ethical approval

The study was approved by the Institutional Review Board (IRB) of Bournemouth University in the UK (No. 42163) and Hamad Bin Khalifa University, Qatar (No. QBRI-IRB-2024-5), and all research was performed in accordance with the relevant guidelines and regulations.

Informed consent

All participants gave their informed consent for participating in the online survey. The Consent Form is available at the Open Science Framework link (https://osf.io/cys8j/).

Additional information

Correspondence and requests for materials should be addressed to Raian Ali.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.