International Journal of Human–Computer Interaction

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/hihc20

Explainability as a Psychological Inoculation: Building Resistance to Digital Persuasion in Online Gambling through Explainable Interfaces

Deniz Cemiloglu, Selin Gurgun, Emily Arden-Close, Nan Jiang & Raian Ali

To cite this article: Deniz Cemiloglu, Selin Gurgun, Emily Arden-Close, Nan Jiang & Raian Ali (2024) Explainability as a Psychological Inoculation: Building Resistance to Digital Persuasion in Online Gambling through Explainable Interfaces, International Journal of Human–Computer Interaction, 40:23, 8378-8396, DOI: 10.1080/10447318.2023.2281744

To link to this article: https://doi.org/10.1080/10447318.2023.2281744

INTERNATIONAL JOURNAL OF HUMAN–COMPUTER INTERACTION
2024, VOL. 40, NO. 23, 8378–8396
https://doi.org/10.1080/10447318.2023.2281744

Explainability as a Psychological Inoculation: Building Resistance to Digital Persuasion in Online Gambling through Explainable Interfaces

Deniz Cemiloglu[a], Selin Gurgun[a], Emily Arden-Close[a], Nan Jiang[a], and Raian Ali[b]

[a]Faculty of Science and Technology, Bournemouth University, UK; [b]College of Science and Engineering, Hamad Bin Khalifa University, Qatar

ABSTRACT

Persuasive interfaces raise ethical concerns when users are unaware of persuasion or find it hard to resist it. Inoculation Theory suggests that attitudes can be inoculated against persuasive attacks. Studies show that disclosure statements in native advertising help people recognize persuasive intent. Likewise, just-in-time disclosure statements in persuasive interfaces may have a similar effect. In this article, *explainable persuasion* was used as an inoculation intervention to build resistance against persuasive interfaces. The effectiveness of this approach was assessed via a 4x2 online experiment, taking online gambling as an illustrative domain. 240 participants (age range 18–73 years, 138 male, 100 female, 2 participants choose not to disclose) were recruited from the UK. Inoculation was delivered through an animated video, while *explainable persuasion* was operationalized through the disclosure of persuasive intent. The findings showed that *explainable persuasion* increased awareness of the presence and risks of persuasive interfaces and strengthened user resistance to persuasive attempts. *Explainable persuasion*, being information-based, can be a cost-effective strategy for helping people stay in control over their digital usage while engaging with persuasive technologies.

KEYWORDS

Explainability; inoculation theory; psychological inoculation; persuasive systems; ethical design; informed consent; online gambling

1. Introduction

With the wealth of information available from various sources such as social media, online advertising, and other digital platforms, the global economy has shifted significantly towards the attention economy, in which businesses compete for people's attention to sell goods and services (Goldhaber, 1997). Interactive online platforms started to employ immersive and persuasive techniques to enrich user interfaces, to engage users and increase business profit (Hogan, 2001). For example, online platforms use persuasive design techniques such as personalized content, notifications, rewards, and social influence to engage users and increase revenue (Fogg, 2003; Spagnolli et al., 2016).

While persuasive interfaces are typically employed to enhance the user experience, their use, in certain cases, may raise ethical concerns. Users may be unaware they are being persuaded (Atkinson, 2006; Smids, 2012), unaware of the unintended negative repercussions of interacting with persuasive interfaces (Berdichevsky & Neuenschwander, 1999), or may find it difficult to resist persuasion. People typically have some knowledge about traditional forms of persuasion, such as those used in physical world advertising and marketing, however, their knowledge of digital persuasive interfaces could be limited (de Pelsmacker & Neijens, 2012). This can hinder the user's ability both to evaluate the persuasion attempt and to reflect and direct their behavior (Timmer et al., 2015). This is more likely to be the case when persuasion is not user initiated but designed to influence in order to benefit a third party (Spahn, 2012). Persuasive interfaces intended to maximize user engagement may also induce or accelerate psychological and cognitive mechanisms related to addictive behavior (Ali et al., 2015; Alrobai et al., 2014; Kuonanoja & Oinas-Kukkonen, 2018). For example, the use of reduction techniques (i.e., reducing the effort to take an action), such as autopay within digital platforms, may speed up users' decision-making process (Cemiloglu, Naiseh, et al., 2021). In such cases, users may rely on intuitive processing, making them prone to cognitive biases (Kahneman, 2011). Concerns regarding system persuasion may increase when the persuasion target is an emotionally or cognitively vulnerable group (Davis, 2009).

To date, various approaches have been taken to explore ethics in persuasive technology. Notably, transparency and user voluntariness have been identified as pivotal elements for establishing ethical persuasive interfaces (Atkinson, 2006; Barral et al., 2014; Timmer et al., 2015). Nonetheless, the practical implementation of transparent persuasive technology has predominantly remained a theoretical concept within academic discourse (Atkinson, 2006; Barral et al.,

CONTACT Deniz Cemiloglu dcemiloglu@bournemouth.ac.uk Faculty of Science and Technology, Bournemouth University, UK

Supplemental data for this article can be accessed online at https://doi.org/10.1080/10447318.2023.2281744.

2014; Smids, 2012; Timmer et al., 2015). Cemiloglu et al. (2023) introduced the concept of *explainable persuasion* as a solution to increase awareness of persuasion attempts and encourage user consent and choice when interacting with persuasive technology. Accordingly, the concept of *explainable persuasion* is defined as: *"The system's transparency about its persuasion attempts so that users can choose to be conscious of how the design may alter their attention or behavior towards certain content or actions and can consent to be subject to it"* (Cemiloglu, Catania, et al., 2021, P378).

In determining the content of *explainable persuasion*, Cemiloglu et al. (2023) give reference to the Informed Consent Theory defined in bioethics literature (Faden & Beauchamp, 1986) and the Persuasion Knowledge Model defined in the consumer research literature (Friestad & Wright, 1994). Accordingly, the content of *explainable persuasion* is suggested to consist of i) information about the persuaders' intention, ii) information about the persuaders' tactics, iii) information about psychological mediators that the persuaders use (i.e., why the tactic is persuasive, what influences the person mentally), and iv) the potential consequences of interacting with such persuasion techniques. Cemiloglu et al. (2023) suggested that by providing such information, *explainable persuasion* may increase awareness of the presence and risks of persuasive interfaces utilized within online platforms and empower the user to make informed decisions. In this regard, *explainable persuasion* could increase awareness of persuasion attempts and foster resistance to persuasion.

In this article, we propose the use of *explainable persuasion* as an inoculation intervention to build resistance against persuasive interfaces. *Explainable persuasion* could be a promising solution for tackling concerns associated with system transparency, ethical considerations, and user autonomy. This is especially relevant in the context of persuasive interfaces, where emotions have the potential to influence decision-making significantly (Hinson et al., 2006). The effectiveness of this approach is evaluated through an online study using an experimental design. We take online gambling as an exemplar domain and application and examine the inoculation effect of *explainable persuasion* on the persuasive design technique of in-game rewards (i.e., pop-up online casino bonuses). Online gambling platforms are equipped with comparable persuasive design elements as those found in social media and gaming platforms to increase player engagement. For example, they reward players with casino bonuses, offer rehearsal options with demo games and enhance the ease of gambling with auto-spin functions. Such persuasive design techniques might contribute to excessive time and money spent on gambling and have the potential to trigger problematic gambling (Cemiloglu, Naiseh, et al., 2021; McCormack & Griffiths, 2013). Players may be unaware both of persuasion having taken place and of the negative consequences of interacting with persuasive gambling interfaces. As a result, monitoring and controlling gambling behavior while interacting with persuasive interfaces may become difficult, especially for at-risk or problem gamblers. While no consensus exists on the addictive nature of other online spaces, such as social media or online streaming platforms, gambling disorder is the first recognized behavioral addiction in the DSM-5, under the category of "Substance-Related and Addictive Disorders" (American Psychiatric Association, 2013). Moreover, As outlined in the DSM-5, gambling disorder demonstrates diagnostic and etiological resemblances to internet and gaming addiction (American Psychiatric Association, 2013). Hence, we consider online gambling as an illustrative domain for persuasion that can lead to harm and where inoculation or other precautionary mechanisms shall be developed. The findings of this study have important implications for understanding the potential of *explainable persuasion* in increasing awareness of persuasion attempts and conferring resistance to persuasive interfaces. They also have practical implications for responsible gambling initiatives as *explainable persuasion* could be included in responsible gambling policy and practices to both enable informed choice and promote safer gambling.

The article is structured as follows. In Section 2, we provide a comprehensive overview of the relevant literature regarding the present study. In Section 3, we provide the methodology of the study. In Section 4, we present the results of the study. In Section 5, we provide a general discussion and underline threats to validity and in Section 6, we provide suggestions for future work.

2. Literature review

2.1. Persuasion and persuasive technology

Persuasion is defined as a conscious effort to shape, reinforce, or change the responses of others (Cameron, 2009; Roloff & Miller, 1980). The goal of persuasion is to alter a person's decisions and behaviors without resorting to force or coercion, and ultimately, the decision to change rests with the person being persuaded (Jones & Simons, 2017). Persuasive technology is defined as "any interactive computing system designed to change people's attitudes or behaviors" (Fogg, 2003, P1). In the persuasive technology literature, different terminologies, such as persuasive technology, persuasive systems, and persuasive interfaces, are used to refer to computer systems designed to alter user behaviors. Persuasive systems are usually grouped into two categories: behavioral change support systems (BCSS), in which users utilize technology to modify their behavior or attitude to attain a self-defined goal (Oinas-Kukkonen, 2013), and systems that persuade users for the persuader's gain (Spahn, 2012). Typical examples of BCSS applications are those that promote positive behaviors such as physical activity, personal well-being, and energy savings (Graml et al., 2011; Langrial et al., 2012; Oyebode et al., 2020). The second category includes interactive online platforms that utilize persuasive interfaces to maximize user engagement, such as social networks, gaming, and online gambling platforms. This study addresses persuasive interfaces designed to influence the users for the persuader's gain.

2.2. Resistance to persuasion

Individuals who recognize persuasion attempts can use a variety of strategies to resist and limit the influence that persuasion has on their decision-making and behaviors (Fransen, Smit, et al., 2015; Fransen, Verlegh, et al., 2015; Zuwerink Jacks & Cameron, 2003). Fransen, Smit, et al. (2015) proposed different strategies people use to resist persuasion, such as avoidance strategies (i.e., simply avoiding persuasion attempts), contesting strategies (i.e., actively counterarguing against the message, the source, or the employed persuasion technique), biased-processing strategies (i.e., comprehending the information in a manner that supports their existing views) and empowerment strategies (i.e., declaring current views rather than opposing the persuasive argument). Counterarguing, which is an instance of contesting strategy, is one of the most commonly employed strategies for resisting persuasion (Fransen, Smit, et al., 2015; Zuwerink Jacks & Cameron, 2003). According to this strategy, when people encounter a persuasive argument, they evaluate it in light of previously held beliefs, and if discrepancies are found, the argument is refuted by generating counterarguments (Wright, 1975). Explicitly revealing the intent of the persuasive argument might increase the likelihood of counterarguing (Amazeen, 2021; Amazeen & Wojdynski, 2019; Compton & Ivanov, 2012).

2.3. Inoculating people against persuasive attack

According to Inoculation Theory, which is an extension of contesting strategies, people's attitudes can be inoculated against persuasive attacks in the same manner as the immune system can be inoculated against viral attacks (McGuire, 1961, 1964). McGuire (1964) suggested that exposing someone to a weakened version of a persuasive attack can help them protect their established attitudes against stronger persuasive attacks that may happen in the future. Inoculation interventions are suggested to trigger resistance to persuasion through two main components: threat and refutational pre-emption (McGuire, 1961, 1964). The threat component works on a more affective basis and warns individuals about their vulnerability to future persuasive attacks. The refutational pre-emption component works on a more cognitive basis. This component first raises arguments that may be used in persuasive attacks and then refutes them to help individuals protect their attitudes. This two-sided approach triggers greater resistance than a one-sided message as, through being introduced to the opposing viewpoint, the individual has been offered a basis for challenging the opposite view (Lumsdaine & Janis, 1953). By motivating individuals to protect their established attitudes and by providing content for counterarguments, psychological inoculation helps people critically analyze persuasion attempts and decide whether to be persuaded (McGuire, 1961, 1964; Pfau et al., 1997). If the persuasion attempt is not aligned with the individual's attitudes and personal goals, the individual may use counterarguments as a defense mechanism to resist persuasion (McGuire, 1961, 1964).

Inoculation interventions can be implemented using either a prophylactic approach (i.e., with the aim of preventing attacks on established attitudes) (McGuire, 1964; Pfau et al., 2004) or a therapeutic approach (i.e., with the aim of building resistance to persuasion among individuals with neutral or opposing attitudes) (Compton, 2020; Van der Linden et al., 2020). Furthermore, inoculation interventions are not only effective on argument-specific resistance but also have the potential to inoculate individuals against the very tactics used in persuasive attacks (Roozenbeek & van der Linden, 2019a, 2019b). Inoculation interventions have been conducted in various contexts, such as advertising, political campaigns, social issues and health (Banas & Rains, 2010). Studies successfully conferred resistance to deceptive food advertising (Mason & Miller, 2013), native advertisements (Amazeen, 2021), fake news (Roozenbeek & Van Der Linden, 2019b), legalization of the use of handguns and marijuana (Pfau et al., 2009), and pressures to smoke cigarettes (Pfau et al., 1992) and consume alcohol (Godbold & Pfau, 2000). Inoculation success has been evaluated with print (Parker et al., 2012) video (Godbold & Pfau, 2000), game-based interventions (Roozenbeek & Van Der Linden, 2019b; Van der Linden et al., 2020) and automatized online systems (Gidron et al., 2023; Levy et al., 2019). According to the inoculation literature (Banas & Rains, 2010; Mason & Miller, 2013; McGuire, 1961, 1964; Pfau et al., 2009; Roozenbeek & van der Linden, 2019a), if an inoculation intervention is successful, participants report:

- Greater elicited threat with regards to the persuasive attack.
- Higher issue involvement levels about the attitudinal object after inoculation.
- Less favorable attitudes towards the object of the persuasive attack.
- Less intention to interact with the object of the persuasive attack.
- Less favorable attitudes towards the persuasive attack.
- Higher likelihood to counterargue against the object of the persuasive attack.

Based on the preceding review of literature and rationale, we raise the following research question:

***RQ1*:** Can inoculation intervention confer resistance against persuasive design techniques used in online gambling platforms?

H1: Participants who receive the inoculation intervention will report (a) higher elicited threat, (b) greater issue involvement and (c) more counterarguments. They will also report (d) less favorable attitudes towards online casino bonuses, (e) less favorable attitudes towards persuasive attack, and (f) lower intention to claim online casino bonuses compared to the control condition.

2.4. Inoculation through peripheral cues

Research suggests that the threat component on its own can confer resistance to persuasion (Kiesler & Kiesler, 1964;

Petty & Cacioppo, 1979). Knowles and Linn (2004) suggested that just as persuasion can result from peripheral cues, so can resistance to persuasion. Inoculation can also work heuristically through peripheral cues requiring minimal cognitive effort (Banas & Miller, 2013). Studies conducted in the advertising domain support this claim and argue that native advertising (when a marketer presents paid content in a manner that closely resembles the publisher's original content to leverage the publisher's credibility (Wojdynski & Golan, 2016)) disclosures can act as forewarning which helps the individual recognize the commercial content (Amazeen, 2021; Amazeen & Vargo, 2021; Amazeen & Wojdynski, 2019; Wojdynski & Golan, 2016). When a persuasive incentive is observed, persuasion knowledge, which consists of information relating to the persuader and the persuasion target, is activated (Friestad & Wright, 1994). The Persuasion Knowledge Model postulates that when individuals have information on both the persuader and the persuasion target (i.e., self), they can analyze the persuasion attempt critically, reducing their susceptibility to persuasion (Livingstone & Helsper, 2006; Panic et al., 2013). Thus, recognizing the persuasive intent helps the individual evaluate the persuasion attempt and resist persuasion if it is not in line with their personal goals (Friestad & Wright, 1994; McGuire, 1964). In this regard, like the use of disclosure statements in native advertising, such as "this celebrity has been paid to appear in this advert," *explainable persuasion* in the form of a just-in-time disclosure statement can potentially inoculate the viewer promoting resistance to persuasion when interacting with persuasive interfaces. Based on the literature review and rationale, we raise the following research question:

***RQ2*:** Can *explainable persuasion* be employed as an inoculation intervention to confer resistance against persuasive design techniques used in online gambling platforms?

While research indicates that threat on its own can confer resistance (Kiesler & Kiesler, 1964; Petty & Cacioppo, 1979), McGuire and Papageorgis (1962) argue that the threat itself is not as impactful as the threat paired with refutational preemption. Moreover, inoculation is suggested to be more effective when delivered multiple times over a specific time period rather than once (Ivanov et al., 2018). Accordingly, we hypothesize:

H2: Participants who receive both the inoculation intervention and *explainable persuasion* during the persuasive attack will report (a) more counterarguments, (b) less favorable attitudes towards online casino bonuses, (c) less favorable attitudes towards the persuasive attack, and (d) lower intention to claim online casino bonuses compared to the control condition.

2.5. Responsible gambling and informed choice

Gambling is recognized as a social and public health issue (Korn & Shaffer, 1999) and in response, governments and gambling providers globally introduced responsible gambling policies and practices to prevent and mitigate the adverse effects of gambling disorder on players and the community (Blaszczynski et al., 2011). The principles of autonomy and informed choice are fundamental to responsible gambling policies and practices (Blaszczynski et al., 2004). Blaszczynski et al. (2011) argued that the main responsibility of the gambling industry is to offer adequate and useful information that will facilitate informed player choices. That is, the gambling industry is obligated to disclose and inform players about games' features and how they work, along with the potential harm and consequences related to interacting with such games. This information should be relevant, accurate, accessible, understandable and provided on a timely basis (Blaszczynski et al., 2013).

Studies have shown that responsible gambling pop-up messages can increase self-awareness among casual gamblers, resulting in more responsible gambling behavior and more informed decisions (Auer & Griffiths, 2016; Monaghan, 2009). However, most published research on the effectiveness of pop-up messages on gambling behavior does not compare problem gamblers to non-problem gamblers (Bjørseth et al., 2020). Caillon et al. (2021) found that informative pop-up messages decreased the illusion of control (i.e., believing that one has control over gambling outcomes) (Cantinotti et al., 2004; Langer, 1975) for at-risk gamblers compared to control participants. Nevertheless, more research is needed to provide insight into the effectiveness of informative messages among different gambler profiles.

Similar to responsible gambling pop-up messages, *explainable persuasion* has the potential to help players assess the implications of interacting with persuasive gambling interfaces so that they can make informed choices. To identify the diverse needs of different players, it is important to evaluate the effectiveness of *explainable persuasion* in conferring resistance to persuasion across different gambler profiles. As no specific direction of the relationship has been proposed previously, we hypothesize that,

H3: There will be a difference in the level of (a) elicited threat, (b) issue involvement, (c) attitudes towards online casino bonuses, (d) intention to claim online casino bonuses, (e) attitudes towards persuasive attack, and (f) number of counterarguments between different problem gambling severity groups.

In addition to exploring differences in study variables among different gambler groups, we also aim to investigate the potential interaction between inoculation condition and problem gambling severity on these study variables.

H4: There will be an interaction between inoculation intervention and problem gambling severity on study variables.

3. Method

3.1. Study design

A 4×2 design was used in the online study. The inoculation intervention was administered through an animated video.

Explainable persuasion was operationalised as a disclosure statement of persuasive intent during the persuasive attack, i.e., a message stating that the casino bonus offer is intended to persuade the player to continue gambling. Inoculation intervention types (inoculation intervention + disclosure of persuasive intent during persuasive attack, inoculation intervention alone, disclosure of persuasive intent during persuasive attack alone, and control) and problem gambling severity (non-problem and low-risk gamblers, moderate-risk gamblers) as determined by the Problem Gambling Severity Index (PGSI) (Ferris & Wynne, 2001a, 2001b), served as the independent variables. Non-problem gamblers and low-risk gamblers were merged into a single group. This group will be referred to as "non-problem + low-risk gamblers." Baseline attitude toward online casino bonuses was used as a covariate. The study design enabled comparison of the influence of inoculation intervention type and problem gambling severity on resistance to persuasion. The dependent variables were elicited threat, issue involvement with responsible gambling, attitudes towards online casino bonuses, intention to claim online casino bonuses, attitudes towards the persuasive attack and number of counterarguments.

Each dependent variable was analyzed independently. The minimum sample size for the study was determined using Statistical Power Analysis (G*Power 3.1.9.2). G*Power is one of the most commonly used sample size calculation methods in the behavioral sciences (Faul et al., 2007). The G-power software predicts the sample size needed for a study based on statistical significance level, effect size, statistical power, and number of predictors (Faul et al., 2007). The study used a 0.05 value of significance level, 0.25 effect size, 80% statistical power, and two predictors: inoculation intervention (4) and problem gambling severity (2) and one covariate, baseline attitude toward online casino bonuses. A 4×2 design required comparing eight groups. For the interaction effect, the numerator *df* was calculated as (4-1) * (2-1) = 3. Accordingly, the total needed sample size by the software was estimated to be 179. This equaled to approximately 23 participants in each condition. The final sample size satisfied the requirements for adequate statistical power. There were 240 participants, with 60 in each of the four conditions.

3.2. Participants

Overall, 240 participants (age range 18–73 years, 138 male, 100 female, 2 participants choose not to disclose) were recruited to the online study through Prolific™ (www.prolific.co), an established online research participant recruitment platform. Participants who had bet daily or weekly on online slot and roulette games in the previous 12 months, who were aged 18 years or older and fluent English speakers were recruited for the study. Participants were informed that the study's objective was to examine their attitudes to casino bonuses used by online gambling websites. Initially, 394 participants were screened. To maintain the study's relevance to digital usage, we excluded problem gamblers from our research. This decision was based on the consideration that their characteristics might not correspond with those of the average user who interacts with persuasive interfaces. Accordingly, individuals who were undergoing treatment or who were experiencing any negative consequences resulting from their gambling were excluded. There were three screening steps for participant recruitment.

1. In the invitation letter, participants were informed that the study was intended for moderate gamblers (i.e., gambling within reasonable and proper limits) and those who thought they may need support were directed to relevant support services.
2. Before participants could take part in the study, they were required to check a box stating they were not experiencing problems due to gambling in the participant information sheet.
3. To avoid recruiting participants who might be unaware of their problems, participants were assessed by the Problem Gambling Severity Index (PGSI) (Ferris & Wynne, 2001a, 2001b), which is a valid and reliable instrument commonly used in gambling research to screen for problem gambling. Participants with a PGSI score of eight or higher were classified as problem gamblers and disqualified from the study. A message through Prolific was sent to disqualified individuals, informing them of where they may receive help. These screening procedures reduced the likelihood of participants experiencing psychological stress or anxiety due to gambling.

3.3. Measures

3.3.1. Problem gambling severity index

The 9-item PGSI was used to assess problem gambling severity (Ferris & Wynne, 2001a, 2001b). The scale includes items related to gambling behavior (e.g., How often have you bet more than you could really afford to lose?) and experienced adverse consequences due to gambling (e.g., How often has your gambling caused any financial problems for you or your household?). Each item is rated on a 4-point scale: 0 never; 1 = sometimes; 2 = most of the time; 3 = almost always. The standard cut-points are 0 = non-problem gambler; 1–2 = low-risk gambler; 3–7 = moderate-risk gambler; and eight or more = problem gambler. Utilizing the PGSI ensured that each condition had an equal number of gambler profiles. The PGSI has high internal consistency and test-retest reliability (Devlin & Walton, 2012; Lopez-Gonzalez et al., 2018; So et al., 2019). Cronbach's alpha was 0.70, indicating good reliability.

3.3.2. Elicited threat

Elicited threat was assessed at Phase 2 using five bipolar adjective pairs (Mason & Miller, 2013; Pfau et al., 1992): unintimidating–intimidating, nonthreatening– threatening, not risky–risky, not harmful–harmful, and safe–dangerous. Participants were given the following scenario:

> Imagine that you are at the end of your gambling session for the day and are ready to leave the gambling website. You receive a notification offering you an extra £20 bonus to spend on a new game if you deposit £20. This notification intends to cause you to rethink your decision of leaving the gambling website. We want to know how this would make you feel.

After reading the scenario, participants were asked to indicate how this would make them feel on a 7-point scale for each adjective pair. The pairs were rated 1 (e.g., unintimidating) to 7 (e.g., intimidating). Greater elicited threat was reflected by higher scores. The reliability of the elicited threat scale was 0.85 ($n = 240$), as assessed by Cronbach's coefficient alpha.

3.3.3. Issue involvement with responsible gambling

Issue involvement was assessed at Phases 1 and 2 using a shortened version of Zaichkowsky's (1985) Personal Involvement Inventory (PII). As in other inoculation studies (Ivanov et al., 2009; Parker et al., 2012), seven items were utilized for the assessment: unimportant–important, irrelevant–relevant, nonessential–essential, of no concern–of concern to me, does not matter–matters to me, useless–useful, and trivial–fundamental. Participants were asked to indicate what responsible gambling meant to them using a 7-point scale for each item. The reliability ratings for the issue involvement scale were Phase 1: 0.86 and Phase 2: 0.82 ($n = 240$), as assessed by Cronbach's coefficient alpha.

3.3.4. Attitudes towards the use of online casino bonuses and the persuasive attack

Attitudes towards online casino bonuses was assessed at Phases 1, 2 and 3 using six bipolar adjective pairs (Pfau, Szabo, et al., 2001; Pfau et al., 2006; Pfau & Burgoon, 1988): foolish–wise, unacceptable–acceptable, wrong–right, unfavorable–favorable, bad–good, and negative–positive on a 7-point scale. The reliability ratings of the attitude scale were Phase 1: 0.93, Phase 2: 0.95, and Phase 3: 0.95 ($n = 240$). Attitude towards the persuasive attack was assessed with the same measure. Reliability (Cronbach's alpha) of the attitude scale for the persuasive attack was 0.85.

3.3.5. Intention to claim online casino bonuses

Intention to claim online casino bonuses was assessed at Phases 1, 2 and 3 using a single item, 0–100-point scale (Compton & Pfau, 2004; Pfau, Park, et al., 2001). The question asked, "on a scale from 0 (no probability) to 100 (certain probability), what is the likelihood you will claim online casino bonuses (e.g., cash bonuses and free spins)?"

3.3.6. Counterarguments

Counterarguments were assessed using a thought-listing technique (Cacioppo et al., 1997; Cacioppo & Petty, 1981). The method used by Reynolds-Tylus et al. (2019) was adopted in the present study. After viewing the pop-up online casino message, participants were instructed to take 90 seconds to list all the thoughts that came to their minds while they viewed the message. Participants were provided with 10 text boxes and were asked to write down each thought in a different box. On the following page, participants were asked to indicate whether each thought was about the content of the pop-up message or not to assess relevance. On the next page, participants were asked to indicate for each thought whether it was unfavorable (i.e., a negative thought about the pop-up message), neutral (i.e., neither favorable nor unfavorable thought about the pop-up message) or favorable (i.e., a positive thought about the pop-up message) to assess valence. Only relevant and negative thoughts were counted as counterarguments, yielding a single metric to assess counter-argumentation. The coding for study variables was verified by another member of the research team.

3.4. Procedure

The study was designed on Qualtrics (https://www.qualtrics.com/), a web-based survey platform. Bournemouth University Research Ethics Committee approved the study on 11 May 2022 (ID: 39653). Data collection began on 5 September 2022 and closed on 9 December 2022. The study consisted of three phases. Flow of participants through the study is shown in Figure 1.

3.4.1. Pre-screening

Participants were screened, and those who met the inclusion criteria were recruited for the study. Participants who were experiencing negative consequences resulting from their gambling and participants with a PGSI score of eight or higher were disqualified from the study. Because problem gambling severity was used as an independent variable, the researcher aimed to enroll an equal number of non-problem + low-risk gamblers and moderate-risk gamblers in the study. Due to the random nature of PGSI scores, 394 participants were screened for the study to achieve this objective. Eventually, 120 non-problem + low-risk gamblers and 120 moderate-risk gamblers were recruited to the study, totaling 240 participants.

3.4.2. Phase 1

In the first phase, participants were asked to provide information about their gambling experience (e.g., number of online gambling accounts and time spent gambling per week). Participants' attitudes towards online casino bonuses, intention to claim online casino bonuses and issue involvement with responsible gambling at baseline were assessed by a questionnaire.

3.4.3. Phase 2

Phase 2 took place one week after Phase 1. In Phase 2, the inoculation intervention video was shown to the participants. 120 participants were assigned to the inoculation intervention condition and 120 participants to the control condition. A matched pair approach was taken since

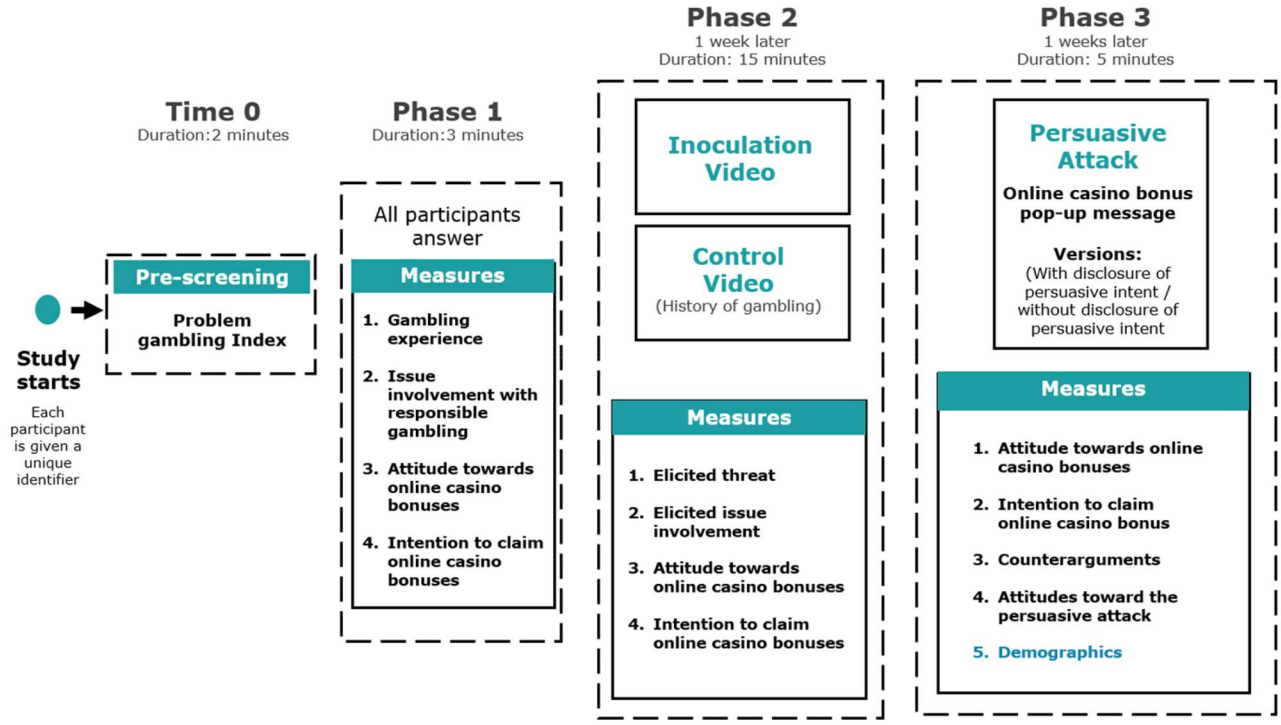

Figure 1. Inoculation study flow.

problem gambling severity was used as an independent variable. Participants in the inoculation and control groups were paired according to their PGSI groups to lessen the impact of confounding factors on the results. Accordingly, an equal number of non-problem + low-risk gamblers and moderate-risk gamblers were randomly allocated to one of the two conditions. In both the inoculation intervention condition and the control condition, there were 60 non-problem- + low-risk gamblers and 60 moderate-risk gamblers, totaling 240 participants.

Participants in the inoculation intervention condition were initially asked how well they knew how online gambling websites can motivate them to gamble on a scale from 0 (no knowledge) to 100 (high knowledge) to elicit threat. Later, participants in this condition watched a 5-minute inoculation video about online casino bonuses. The inoculation video contained arguments that gambling operators may use to persuade players to claim online casino bonuses and refutations of these arguments. Participants were then asked to confirm that they had watched the video. As an attention check, they were asked two multiple-choice and one open-ended question about the video.

Participants in the control condition watched a 5-minute video about the history of gambling. The control video presented information about gambling from ancient times to the digital age. Participants were then asked to confirm that they had watched the video. As an attention check, participants in this condition were also asked two multiple-choice and one open-ended question about the video. Following the videos, participants across all conditions were asked to answer a questionnaire on elicited threat, issue involvement with responsible gambling, attitudes towards online casino bonuses and intention to claim online casino bonuses.

3.4.4. Phase 3

Phase 3 took place one week after Phase 2. McGuire (1964) suggested that a delay is needed between the inoculation intervention and the attack as it takes time to counterargue and generate arguments for defense (McGuire, 1964). In Phase 3, both the inoculation intervention condition and the control condition received the following scenario:

> Imagine you have been gambling at a gambling website called Fun & Bet Casino. You realize that you lost more money than you expected in your gambling session and are considering leaving the website. Just before you close the website, a pop-up message appears.

After reading the scenario, the persuasive attack was presented in the form of a pop-up online casino bonus message resembling those used in gambling websites (See Appendix, supplementary material). Half the participants in the inoculation intervention condition and the control condition were exposed to the pop-up message with a threat forewarning in the form of a disclosure statement about the persuasive intent of the pop-up message (i.e., *explainable persuasion*). The other half of the participants in the inoculation intervention condition and the control condition were exposed to the same pop-up message without the disclosure statement. Following the persuasive attack, all participants answered a questionnaire on counter-argumentation, attitudes towards online casino bonuses, attitudes towards the persuasive attack (i.e., pop-up bonus offer), and intention to claim online casino bonuses. Participants who were exposed to the pop-up message with a disclosure statement of persuasive intent were asked how likely they were to click the "learn more" button to find out how persuasive features may impact their gambling behavior with a 5-point scale (1 = very unlikely, and 5 = very likely). Participants were

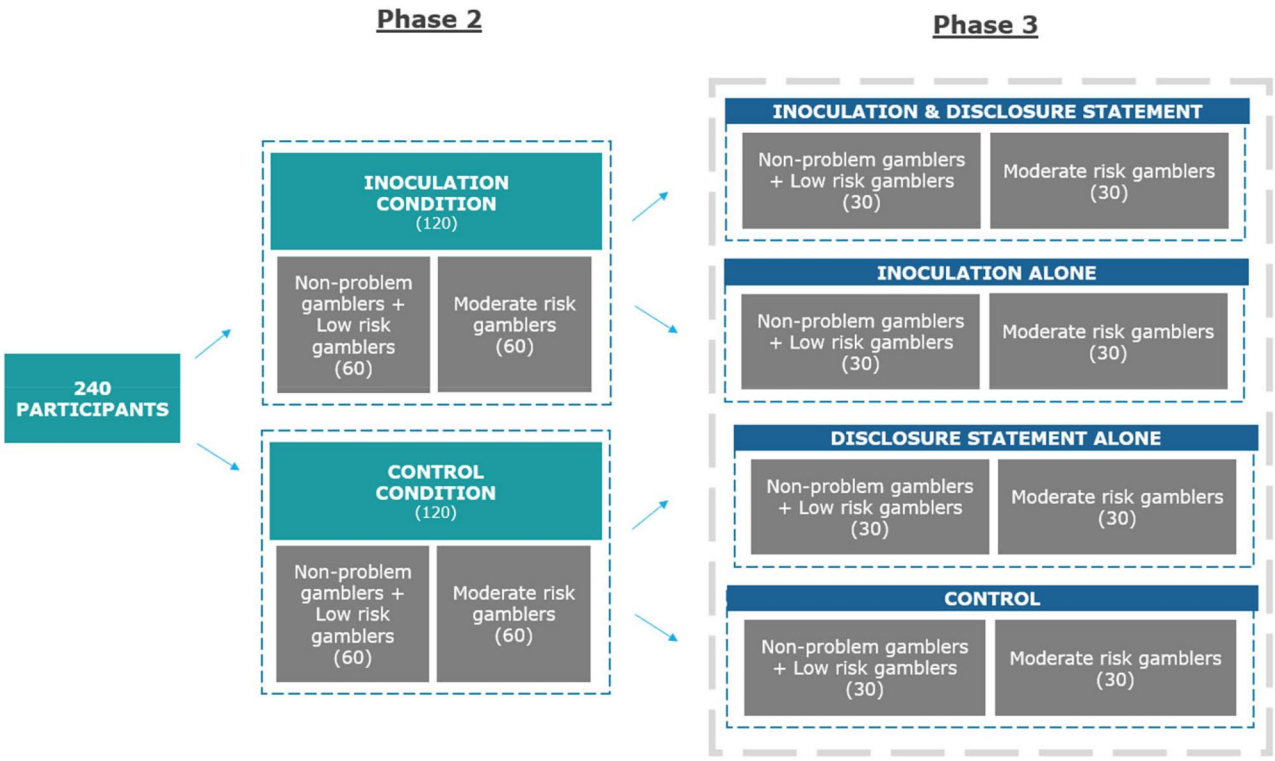

Figure 2. Inoculation study experimental conditions.

also asked to give a rationale for their answers. In the last phase, demographic information was collected from participants. Participants who completed all three phases of the study received £2.70 for their participation. Four participants who did not provide sensible answers were excluded from the study. The flow of participants through the intervention is detailed in Figure 2.

3.5. Materials

Materials used include the inoculation intervention video, the control video, and the persuasive attack.

3.5.1. Inoculation intervention video

The script of the inoculation intervention video consisted of three parts. The first part of the script was intended to induce threat. Participants were warned that while many players control their gambling and enjoy it as a leisure activity, gambling operators successfully create online casino bonuses such as cash bonuses or free spins to persuade players to spend more time and money than they initially intended. The second and third parts consisted of arguments that gambling operators may use to persuade players to claim online casino bonuses and refutations of these arguments. The arguments for claiming online casino bonuses included: (i) getting a head start by spending less of your own money and (ii) trying out exciting new games for free through exclusive bonuses. These arguments reflected those used in online casino bonus advertisements. The arguments against claiming online casino bonuses included: (i) online casino bonuses being subject to specific play requirements and the use of words like "bonus" and "free" reducing the apparent cost of play requirements, (ii) online casino bonuses disrupting players from their responsible gambling goals by acting as triggers and making it difficult for players to reflect on future repercussions. These arguments were based on the findings of a studies that examined the relationship between persuasive interfaces and addictive behavior (Cemiloglu, Naiseh, et al., 2021; McCormack & Griffiths, 2013). In total, the inoculation video script was 417 words. The video script was animated with PowToon (https://powtoon.com), a web-based animation platform. The text was narrated by a British-accented female narrator (See Figure 3). See Appendix for the video text (supplementary material).

3.5.2. Control video

The script of the control video consisted of six parts. The script gave a review of gambling throughout history and consisted of the earliest foundations, the Ancient World, the Middle Ages, the Enlightenment, Modern History, and the Digital Age. The content of the script was based on information presented in online articles (Encyclopædia Britannica, 1998; Reader's Digest, 2020). In total, the inoculation video script was 396 words. Similar to the inoculation intervention video, the video script was animated with PowToon and narrated by the same British-accented female narrator (See Figure 4). See Appendix for the video text (supplementary material).

3.5.3. Persuasive attack

The persuasive attack was in the form of a pop-up casino bonus offer for a new online slot game resembling those

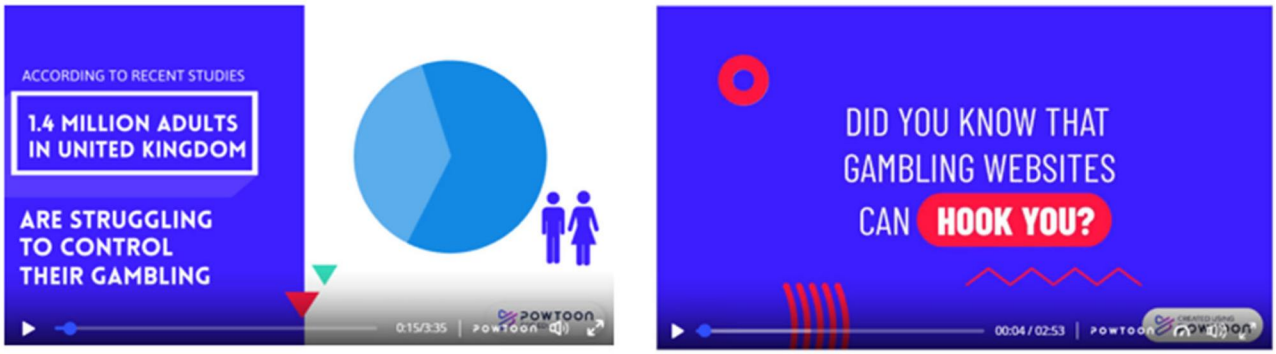

Figure 3. SCREENSHOTS of the inoculation video.

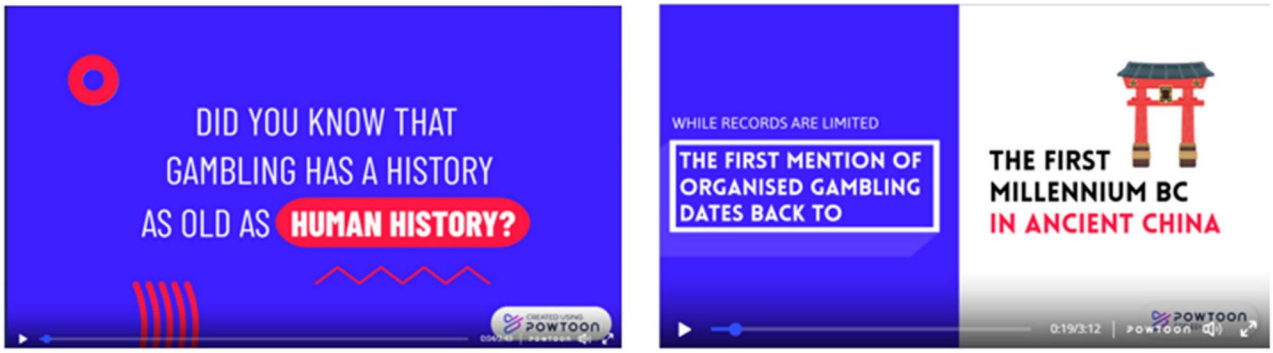

Figure 4. SCREENSHOTS of the control video.

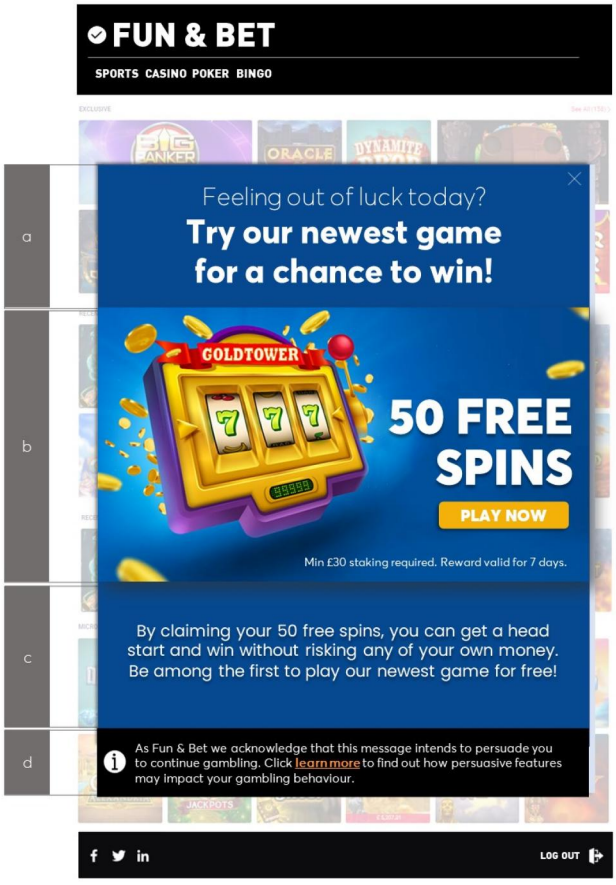

Figure 5. Persuasive attack with disclosure statement.

used in gambling websites (See Figure 5). The choice of a pop-up casino bonus as the persuasive attack was made due to its prevalence across nearly all online gambling platforms, making it a representative example of persuasive design techniques. The study took place within a controlled environment, ensuring that the influence of the flow experience during gambling remained regulated. This allowed for a singular focus on determining whether an individual would accept the offer. The similarity, validity, and clarity of the pop-up casino bonus offer was evaluated by two responsible gambling officials, four academics, and one ex-problem gambler.

We selected a slot game as research indicates that slot machine gamblers are more susceptible to irrational thinking and biases than players of other games (Walker, 1992). Furthermore, the short period between betting and the outcome of such games may result in less self-aware betting (Monaghan & Blaszczynski, 2010).

The pop-up message consisted of three parts. The top part addressed the player with, “Feeling out of luck today? Try our newest game for a chance to win!” The middle part introduced a new game called Gold Tower with a colorful visual and offered 50 free spins. Similar to typical online casino bonus offers, the fine print detailed play requirements. The fine print read, “Min £30 staking required. Reward valid for 7 days.” The pop-up message had a clickable button that was labelled with the call-to-action phrase "Play Now." The bottom part further advertised the benefits of claiming the offer. Two different versions of the pop-up message were utilized in the study. One version included a

disclosure statement about the persuasive intent of the pop-up message in the footer, while the other version did not (See Appendix, supplementary material). All other aspects were identical in both versions. The disclosure statement was as following:

> As Fun & Bet, we acknowledge that this message intends to persuade you to continue gambling. Click **learn more** to find out how persuasive features may impact your gambling behavior.

3.6. Data analysis

Data was analyzed using SPSS version 28. Non-parametric tests were used when appropriate, as the data was not normally distributed. Three main analyses were performed.

Analysis of Covariance (ANCOVA) was used as the main analysis to test the effect of the inoculation intervention and problem gambling severity on elicited threat, issue involvement, attitudes towards online casino bonuses, intention to claim online casino bonuses, counterarguments, and attitudes towards the persuasive attack. Baseline attitude toward online casino bonuses was used as a covariate in all analyses. 2×2 ANCOVA were used to test three dependent variables at Phase 2. The dependent variables were elicited threat, attitudes towards online casino bonuses, intention to claim online casino bonuses and issue involvement with responsible gambling measured at Phase 2. Inoculation condition (inoculation, no inoculation) and problem gambling severity (non-problem + low-risk gamblers, moderate-risk gamblers) served as the independent variables. A 4×2 ANCOVA was used to test four dependent variables at Phase 3. The dependent variables were attitudes towards online casino bonuses, intention to claim online casino bonuses, attitudes towards persuasive attack and counterarguments measured at Phase 3. Inoculation condition (inoculation and disclosure, inoculation only, disclosure only, control) and problem gambling severity (non-problem + low-risk gamblers, moderate-risk gamblers) served as the independent variables.

Spearman correlations were used to analyze the association between continuous and ordinal variables (Sheskin, 2003). Data from the open-ended question was analyzed using thematic analysis (Braun & Clarke, 2006). The coding was verified by another member of the research team.

4. Results

4.1. Participant demographics

In total, 240 participants completed the online study. Nine participants reported that they work or have worked in the gambling industry. Table 1 summarizes demographics.

Table 1. Participant demographics.

N	240
Age: M (SD)	38.3 (11.1)
Age: Range	18–73
Gender: Males (%)	138 (57.5)
Females (%)	100 (41.7)
Prefer not to say (%)	2 (0.8)
Gambling Activity Days Per Week: M (SD)	2.85 (1.96)
Number of Online Gambling Accounts (%)	
1 account	12.9
2 accounts	20.0
3 accounts	21.7
4 accounts	10.0
5 accounts	5.0
6 or more accounts	30.4
Problem Gambling Severity Index (%)	
Non-problem gambler	26.7
Low-risk gambler	23.3
Moderate-risk gambler	50.0
Education (%)	
Compulsory school education completed	15.4
Vocational training	9.2
College	23.8
University degree	38.8
Postgraduate qualification (e.g., MSc, PhD)	12.9
Employment (%)	
Full-time employment	55.8
Part-time employment	15.4
Self-employed	7.9
Unemployed	8.3
Student	2.1
Retired	2.5
Homemaker	7.1
Other	0.8

4.2. Manipulation check: Elicited threat

Researchers have suggested that threat vulnerability is required for inoculation to work (Godbold & Pfau, 2000; McGuier, 1962). A two-way ANCOVA was run to examine the effects of the inoculation intervention and problem gambling severity on elicited threat at Phase 2 after controlling for baseline attitudes towards online casino bonuses. There was no statistically significant interaction between problem gambling severity and experimental condition on elicited threat levels in Phase 2, $F(1, 235) = 0.1$, NS. Therefore, the main effects of problem gambling severity and inoculation intervention were analyzed.

There was a statistically significant main effect of the inoculation intervention on Phase 2 elicited threat levels, $F(1, 235) = 4.7$, $p = 0.03$, partial $\eta 2 = 0.02$. The adjusted marginal mean of elicited threat level for the inoculation condition ($M = 4.7$, $SE = 0.1$) was higher than the no inoculation condition ($M = 4.4$, $SE = 0.1$), a statistically significant difference of 0.3 in mean scores (95% CI, 0.3 to 0.6). There was no statistically significant main effect of problem gambling severity on Phase 2 elicited threat levels, $F(1, 235) = 3.1$, $p = 0.08$, partial $\eta 2 = 0.01$. Due to the statistically significant main effect of the inoculation intervention on Phase 2 elicited threat levels, it was considered acceptable to test inoculation theory.

For the inoculation intervention condition, a Spearman's rank-order correlation analysis revealed a statistically significant negative correlation between elicited threat and Phase 2 attitudes towards online casino bonuses, $r_s(238) = -0.2$, $p < 0.01$. That is, participants who had lower elicited threat scores were more likely to have more positive attitudes towards online casino bonuses and vice-versa. The correlation matrix for the study variables is shown in Appendix (supplementary material).

Table 2. P2 Issue involvement with responsible gambling.

	Intervention groups			
	Non + Low-risk gamblers		Moderate-risk gambler	
P2 Issue Involvement	Inoculation	No Inoculation	Inoculation	No Inoculation
M	6.4	6.5	6.5	6.1
(SD)	0.6	0.6	0.5	0.7
M(adj)	6.4	6.5	6.5	6.2
(SE)	0.08	0.08	0.08	0.08

4.3. Issue involvement with responsible gambling

A two-way ANCOVA was run to examine the effects of the inoculation intervention and problem gambling severity on issue involvement with responsible gambling at Phase 2 after controlling for baseline attitudes towards online casino bonuses. There was a statistically significant interaction between problem gambling severity and experimental condition on Phase 2 issue involvement with responsible gambling, F(1, 235) = 6.6, $p = 0.01$, partial $\eta 2 = 0.02$. Therefore, an analysis of the simple main effects for problem gambling severity and inoculation intervention was performed. Means, adjusted means, standard deviations and standard errors are presented in Table 2.

The effect of inoculation intervention on issue involvement levels at Phase 2 for the non + low-risk gambler group was not statistically significant, $F(1, 235) = 1.0$, NS, whereas the effect of inoculation intervention for the moderate-risk gambler group was statistically significant, $F(1, 235) = 6.8$, $p = 0.01$, partial $\eta 2 = 0.02$. Within the moderate-risk gambler group, the inoculation group had a higher issue involvement level compared to the no inoculation group (95% CI, 0.07 to 0.5).

The effect of problem gambling severity on issue involvement levels at Phase 2 for inoculation condition was not statistically significant, $F(1, 235) = 0.7$, NS, whereas the effect of problem gambling severity for the no inoculation condition was statistically significant, $F(1, 235) = 7.5$, $p = 0.006$, partial $\eta 2 = 0.03$. For the no inoculation condition, the non + low-risk gambler group had a higher issue involvement level compared to the moderate-risk gambler group (95% CI, 0.09 to 0.5).

For the inoculation intervention condition, a Spearman's rank-order correlation analysis revealed a statistically significant positive correlation between baseline issue involvement with responsible gambling and Phase 2 issue involvement with responsible gambling, $r_s(238) = 0.3$, $p < 0.001$. That is, participants who had high issue involvement with responsible gambling at baseline also had high issue involvement with responsible gambling at Phase 2.

4.4. Attitudes towards online casino bonuses

A two-way ANCOVA was run to examine the effects of the inoculation intervention and problem gambling severity on attitudes towards online casino bonuses at Phase 2 after controlling for baseline attitudes towards online casino bonuses. There was no statistically significant interaction between problem gambling severity and experimental condition on Phase 2 attitudes towards online casino bonuses, F(1, 235) = 1.3, NS. Therefore, the main effects of problem gambling severity and inoculation intervention were analyzed.

There was a statistically significant main effect of the inoculation intervention on Phase 2 attitudes towards online casino bonuses, F(1, 235) =24.2, $p < 0.001$, partial $\eta 2 = 0.09$. The adjusted marginal mean of attitudes towards online casino bonuses for the inoculation condition (M = 4.0, SE = 1.1) was lower than the no inoculation condition (M = 4.7, SE = 0.1). Lower scores meant less favorable attitudes towards online casino bonuses. There was no statistically significant main effect of problem gambling severity on attitudes towards online casino bonuses, F(1, 235) =0.4, NS.

A two-way ANCOVA was run to examine the effects of the inoculation intervention type and problem gambling severity on attitudes towards online casino bonuses at Phase 3 after controlling for baseline attitudes towards online casino bonuses. See Table 3. There was no statistically significant interaction between problem gambling severity and type of inoculation intervention on Phase 3 attitudes towards online casino bonuses, F(3, 231) = 1.2, $p = 0.3$, NS. Also, there was no statistically significant main effect of problem gambling severity F(1, 231) = 0.7, NS or type of inoculation intervention, F(3, 231) = 1.1, NS.

4.5. Intention to claim online casino bonuses

A two-way ANCOVA was run to examine the effects of the inoculation intervention and problem gambling severity on intention to claim online casino bonuses at Phase 2 after controlling for baseline attitudes towards online casino bonuses. There was no statistically significant interaction between problem gambling severity and experimental condition on Phase 2 intention to claim online casino bonuses, F(1, 235) = 0.6, NS. Therefore, the main effects of problem gambling severity and inoculation intervention were analyzed.

There was a statistically significant main effect of the inoculation intervention at Phase 2 intention to claim online casino bonuses, F(1, 235) =5.4, $p = 0.02$, partial $\eta 2 = 0.02$. The adjusted marginal mean of intention to claim online casino bonuses for the inoculation condition (M = 54.2 SE = 2.6) was lower than the no inoculation condition (M = 62.9, SE = 2.6), a statistically significant difference of 8.6 in mean scores (95% CI, 1.3 to 16.0). There was no statistically significant main effect of problem gambling severity on intention to claim online casino bonuses, F(1, 235) =0.5, NS.

A two-way ANCOVA was run to examine the effects of inoculation intervention type and problem gambling severity on intention to claim online casino bonuses at Phase 3 after controlling for baseline attitudes towards online casino bonuses. See Table 4. There was no statistically significant interaction between problem gambling severity and type of inoculation intervention on Phase 3 intention to claim online casino bonuses, F(3, 231) = 1.0, NS. Also, there was no statistically significant main effect of problem gambling

Table 3. P3 Attitudes towards online casino bonuses.

	Inoculation conditions			
P3 Attitudes towards Online Casino Bonuses	Inoculation + Disclosure	Inoculation Alone	Disclosure Alone	Control
All Participants (n: 240)				
M	4.0	4.1	4.2	4.3
(SD)	1.3	1.5	1.5	1.5
M(adj)	4.0	4.0	4.2	4.4
(SE)	0.1	0.1	0.1	0.1
Non + Low-risk gamblers (n:120)				
M	4.0	4.2	4.0	4.2
(SD)	1.2	1.4	1.6	1.4
M(adj)	3.9	4.2	4.1	4.1
(SE)	0.2	0.2	0.2	0.2
Moderate-risk gambler (n:120)				
M	4.0	4.0	4.3	4.4
(SD)	1.4	1.5	1.4	1.6
M(adj)	4.0	3.8	4.3	4.7
(SE)	0.2	0.2	0.2	0.2

Table 4. P3 Intention to claim online casino bonuses.

	Inoculation conditions			
P3 Intention to Claim Online Casino Bonuses	Inoculation + Disclosure	Inoculation Alone	Disclosure Alone	Control
All Participants (n: 240)				
M	47.9	53.9	56.0	55.0
(SD)	32.3	31.5	31.1	33.3
M(adj)	47.6	52.8	56.6	55.7
(SE)	4.0	4.0	4.0	4.0
Non + Low-risk gamblers (n:120)				
M	40.8	55.4	54.8	53.6
(SD)	32.7	32.4	30.9	29.3
M(adj)	40.5	54.9	56.2	51.4
(SE)	5.6	5.6	5.6	5.6
Moderate-risk gambler (n:120)				
M	54.9	52.4	57.1	56.3
(SD)	30.9	31.0	31.9	37.3
M(adj)	54.7	50.7	57.1	59.9
(SE)	5.6	5.6	5.6	5.6

Table 5. P3 Attitudes towards persuasive attack.

	Inoculation conditions			
P3 Attitudes towards Persuasive Attack	Inoculation + Disclosure	Inoculation Alone	Disclosure Alone	Control
All Participants (n: 240)				
M	3.7	3.8	4.0	4.1
(SD)	1.4	1.3	1.5	1.6
M(adj)	3.7	3.8	4.0	4.2
(SE)	0.2	0.2	0.2	0.2
Non + Low-risk gamblers (n:120)				
M	3.5	4.0	4.1	3.9
(SD)	1.2	1.4	1.7	1.4
M(adj)	3.5	3.9	4.1	3.8
(SE)	0.2	0.2	0.2	0.2
Moderate-risk gambler (n:120)				
M	4.0	3.7	3.9	4.3
(SD)	1.5	1.3	1.4	1.7
M(adj)	3.9	3.6	3.9	4.6
(SE)	0.2	0.2	0.2	0.2

severity (F (1, 231) = 1.4, NS and type of inoculation intervention, F(3, 231) = 1.0, NS.

4.6. Attitudes towards the persuasive attack

A two-way ANCOVA was run to examine the effects of inoculation intervention type and problem gambling severity on attitudes towards the persuasive attack at Phase 3 after controlling for baseline attitudes towards online casino bonuses. See Table 5. There was a statistically significant interaction between problem gambling severity and experimental condition on attitudes towards the persuasive attack at Phase 3, F(3, 235) = 2.6, $p = 0.04$, partial $\eta 2 = 0.03$. Therefore, an analysis of the simple main effects for problem gambling severity and inoculation intervention was performed.

The effect of inoculation intervention on attitudes towards the persuasive attack at Phase 3 for the non + low-risk gambler group was not statistically significant, $F(3, 231) = 1.4$, NS, whereas the effect of inoculation intervention for

Table 6. P3 Counterarguments.

	Inoculation conditions			
P3 Counterarguments	Inoculation + Disclosure	Inoculation Alone	Disclosure Alone	Control
All Participants (n: 240)				
M	3.4	2.4	2.1	2.5
(SD)	2.6	2.0	1.8	2.0
M(adj)	3.4	2.4	2.0	2.4
(SE)	0.3	0.3	0.3	0.3
Non + Low-risk gamblers (n:120)				
M	3.5	2.7	1.9	2.7
(SD)	2.7	2.1	1.5	2.1
M(adj)	3.5	2.7	1.9	2.7
(SE)	0.4	0.4	0.4	0.4
Moderate-risk gambler (n:120)				
M	3.2	2.0	2.2	2.2
(SD)	2.6	1.8	2.0	1.8
M(adj)	3.2	2.0	2.2	2.1
(SE)	0.4	0.4	0.4	0.4

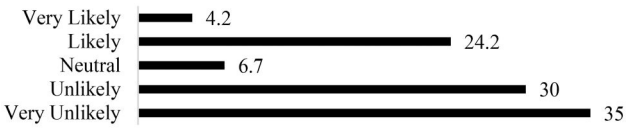

Figure 6. Likelihood to learn more about the disclosure statement.

the moderate-risk gambler group was statistically significant, $F(3, 231) = 2.8$, $p = 0.04$, partial $\eta 2 = 0.03$. Within the moderate-risk gambler group, the inoculation only condition group had less favorable attitudes towards the persuasive attack compared to the control group (i.e., no inoculation and no disclosure of persuasive intent during the attack) with a statistically significant difference of 0.9 in mean scores (95% CI, 0.07 to 1.9).

The effect of problem gambling severity on attitudes towards the persuasive attack at Phase 3 was only significant within the control condition (i.e., no inoculation and no disclosure of persuasive intent during attack), $F(1, 231) = 5.3$, $p = 0.02$, partial $\eta 2 = 0.02$. For the control condition, the non + low-risk gambler group had less favorable attitudes towards the persuasive attack compared to the moderate-risk gambler group, with a statistically significant difference of 0.8 in mean scores (95% CI, 0.1 to 1.4).

4.7. Counterarguments

A two-way ANCOVA was run to examine the effects of inoculation intervention type and problem gambling severity on the number of counterarguments at Phase 3 after controlling for baseline attitudes towards online casino bonuses. Means, adjusted means, standard deviations and standard errors are presented in Table 6. There was no statistically significant interaction between problem gambling severity and experimental condition on the number of counterarguments at Phase 3, $F(3, 231) = 0.7$, NS. Therefore, the main effects of problem gambling severity and inoculation intervention were analyzed.

There was a statistically significant main effect of inoculation intervention type on the number of counterarguments, $F(3, 231) = 4.3$, $p = 0.005$, partial $\eta 2 = 0.05$. The adjusted marginal mean score of counterarguments for the inoculation + disclosure condition was higher than the disclosure only condition, with a statistically significant difference of 1.3 in mean scores (95% CI, 0.3–2.3). That is, participants in the inoculation + disclosure condition generated more counterarguments than participants in the disclosure only condition. There was no statistically significant main effect of problem gambling severity on attitudes towards online casino bonuses, $F(1, 231) = 1.26$, NS.

4.8. Desire to learn about the disclosure statement

As shown in Figure 6, out of 120 participants who were shown the disclosure statement during the persuasive attack, only 28.4% stated that they would want to learn how persuasive features may impact their gambling behavior. There was no significant difference in desire to learn based on gender, PGSI group or inoculation condition at Phase 2.

Participants who were asked how likely they were to click the "learn more" button were also asked to give a rationale for their answers. Figure 7 illustrates the rationale provided by the participants for their decision to either engage or not engage with the disclosure statement (i.e., *explainable persuasion*).

Participants who stated that they would like to learn more about how persuasive features may impact their gambling behavior indicated that providing such explanations will show the integrity of the gambling operators and give players control over their gambling decisions. One participant stated that such information could be especially beneficial when players are chasing losses, as interacting with the information can disrupt such harmful behavior. Participants who said they would not learn more about how persuasive features may impact their gambling behavior indicated they would not be interested in such information. This lack of interest was due to prior knowledge of the persuasive techniques employed by gambling operators, denial of gambling problems, immersion in gambling, desensitization to system warnings in general, and disinterest in the pop-up online casino bonus offer used in the study. Some participants expressed mistrust in gambling operators, claiming that such information will be "superficial" and offered just to comply with regulations. Several participants indicated concerns

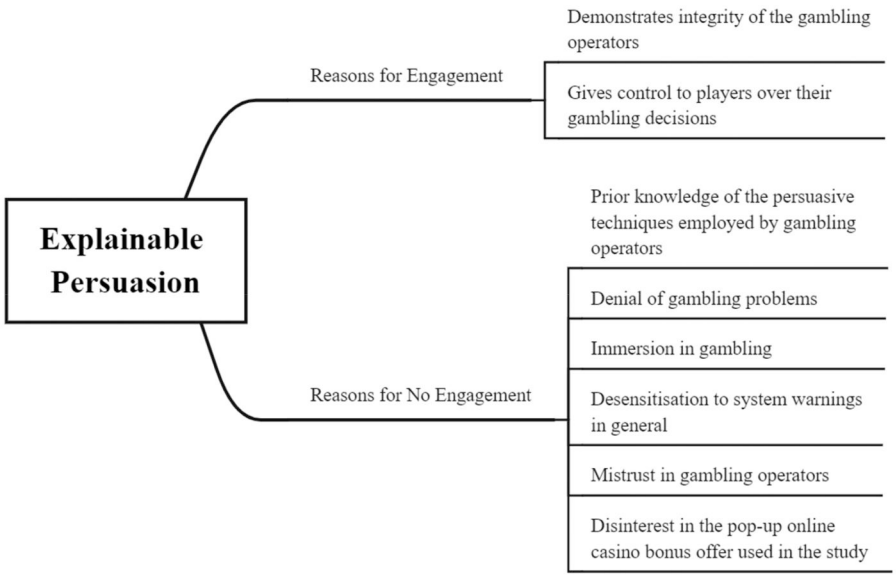

Figure 7. Rationale for engagement with explainable persuasion.

regarding the presentation of the disclosure statement. Participants claimed that the disclosure statement was difficult to read due to fine print and lengthy wording, neither of which encourage responsible gambling behavior.

5. Discussion

The current article proposed the use of *explainable persuasion* as an inoculation intervention and evaluated its effectiveness in building resistance against persuasive design techniques used in online gambling platforms.

At Phase 2, the effect of the inoculation intervention (i.e., inoculation, no inoculation) and problem gambling severity were analyzed. The inoculation intervention effectively reduced positive attitudes towards online casino bonuses and lowered participants' intention to claim online casino bonuses for both problem gambling severity groups. Therefore, Phase 2 findings supported H1. These findings are comparable to inoculation studies in other domains (Compton & Pfau, 2004, 2008).

The inoculation intervention was successful in increasing issue involvement levels with responsible gambling only for moderate-risk gamblers at Phase 2. Within the moderate-risk gambler group, the inoculation condition group reported higher levels of issue involvement compared to the no inoculation group, while no difference was observed within non + low-risk gambler groups. Therefore, Phase 2 findings supported H4. Also, for the no inoculation condition, the non + low-risk gambler group reported higher levels of issue involvement compared to the moderate-risk gambler group. This finding supported H3. This difference may be attributable to participants' pre-existing issue involvement levels. If issue involvement levels are extremely low or high, the inoculation intervention will fail to generate threat since individuals might not worry about their attitudes being attacked or may already have entrenched attitudes (Compton & Pfau, 2009; Pfau et al., 1997). Accordingly, inoculation may have worked better for the moderate-risk gambler group due to their level of issue involvement with responsible gambling, which could be considered optimal for an inoculation intervention. Also, given that the non + low-risk gambler group may already be highly involved with responsible gambling, no change may have been observed following the inoculation intervention. Since the non + low-risk gambler group generally gambles within appropriate levels and may not be concerned with problem gambling (Caillon et al., 2021), they may not have been motivated to process the content of the inoculation video, which may have impacted the results (Amazeen, 2021; Petty & Cacioppo, 2012).

At Phase 3, the effects of the inoculation intervention type (i.e., inoculation intervention + disclosure of persuasive intent during persuasive attack, inoculation intervention alone, disclosure of persuasive intent during persuasive attack alone, and control) and problem gambling severity were analyzed. The study revealed a discernible trend in the data suggesting that participants in inoculation intervention + disclosure condition reported the least positive attitudes towards online casino bonuses and persuasive attack, the least intention to claim online casino bonuses and the highest number of counterarguments against online casino bonuses. This trend was followed by the participants in the inoculation intervention alone condition and disclosure alone condition, respectively. Therefore, Phase 3 findings provided support for H2. There was no statistically significant main effect on attitudes towards online casino bonuses and intention to claim online casino bonuses at Phase 3. This may be due to the time between Phase 2 and Phase 3 (i.e., one week). Even though some researchers suggest that delay could be helpful for inoculation success (McGuire, 1964), Banas and Rains's (2010) a meta-analysis on inoculation research demonstrated that inoculation treatments may lose their effectiveness over time, as motivation to defend attitudes may fade (Insko, 1967). While evaluations for Phase 2 were carried out immediately after the inoculation phase, evaluations for Phase 3 were carried out one week

after the inoculation phase, meaning that the inoculation effect may have diminished.

Regarding attitudes towards the persuasive attack, the findings revealed a statistically significant interaction between problem gambling severity and inoculation intervention type on attitudes towards the persuasive attack. Within the moderate-risk gambler group, the inoculation only condition had less favorable attitudes towards the persuasive attack compared to the control group, while no difference was observed within the non + low-risk gambler groups. Therefore, Phase 3 findings supported H4. This finding is comparable to evidence suggesting that participants who are most susceptible to fake news benefit the most from inoculation interventions (Roozenbeek & van der Linden, 2019a). In other words, the inoculation intervention benefited participants with a greater risk of problem gambling by elevating their negative attitudes towards the persuasive attack. Also, for the control condition (i.e., no inoculation and no disclosure of persuasive intent during attack), the non + low-risk gambler group had less favorable attitudes towards the persuasive attack compared to the moderate-risk gambler group. This finding supported H3.

Regarding counterarguments, the findings revealed a statistically significant main effect of inoculation intervention type on the number of counterarguments. The number of counterarguments for the inoculation + disclosure condition was higher than the disclosure only condition. These findings showed that *explainable persuasion* has the potential to build resistance against persuasive interfaces when coupled with prior inoculation intervention. Compton (2013) indicated that when the inoculation effect diminishes, booster doses may be used to maintain immunity against persuasive attacks. In this light, it is possible that *explainable persuasion* functioned as an inoculation booster dose. This finding also provides support to the argument that threat itself is not as impactful as the threat paired with refutational pre-emption (McGuire & Papageorgis, 1962). Regarding practical applications, similar to multimedia tools used for cybersecurity awareness and education (Albayram et al., 2017; Zhang-Kennedy & Chiasson, 2022), short films and animations, digital games, comics and learning modules could be utilized for inoculation, and *explainable persuasion* could be utilized as a booster dose to sustain the inoculation effect. Such an application may also function as a proactive measure to reduce the habituation effect that may occur with repeated exposure to inoculation content. Habituation happens when a user becomes less responsive to stimuli after repeated exposure (Kim & Wogalter, 2009). Thus, by appearing as salient stimuli, *explainable persuasion* may mitigate the negative effect of habituation.

In terms of the likelihood of engaging with disclosure statements during persuasive attacks, only 28.4% of participants reported wanting to learn how persuasive features may impact their gambling. This lack of interest was attributable to prior knowledge of the persuasive techniques employed by gambling operators, denial of gambling issues, immersion in gambling, desensitization to system warnings in general, disinterest in the pop-up online casino bonus offer used in the study and mistrust in gambling operators. Similar findings were reported by Cemiloglu et al. (2023). Future research could examine how to deliver *explainable persuasion* based on these findings. For example, one approach to address desensitization to system warnings, in general, could be presenting *explainable persuasion* in different formats over time, such as changing the layout or wording, as this can facilitate attention switch and maintenance (Kim & Wogalter, 2009). Moreover, another approach to address denial and mistrust could be related to explanation framing. Positive framing (i.e., emphasizing the benefits of reducing interaction with persuasive interfaces), as opposed to negative framing (i.e., emphasizing the negative consequences of interacting with persuasive interfaces), could help address this negative perception.

This study has a number of limitations. In terms of internal validity, one consideration is social desirability bias. Participants' reported base issue involvement level with responsible gambling was 6.3 (SD = 0.7), with seven being the highest value. Some participants may have been dishonest about their involvement with responsible gambling in order to appear in a favorable light. This may have hindered the ability to observe the change in issue involvement after inoculation intervention. Another consideration is related to the utilization of the persuasive design technique in-game rewards (i.e., pop-up online casino bonus) for the inoculation study. As Cemiloglu et al. (2023) report, in a free recall setting, the most recalled persuasive design technique was in-game rewards (74.4%). Since participants were aware of the use of in-game rewards, the inoculation intervention might have been less effective since players had already applied contesting strategies against rewards. Future research might benefit from utilizing persuasive design techniques that are less well-recognized, such as self-monitoring or social norms.

Considering ecological validity, one consideration is the pop-up online casino bonus used in the study. It is possible that the graphical design or the offer of the pop-up online casino bonus did not resemble those used in online gambling platforms, and as a result, the participants did not find it to be realistic. However, to address this issue, the pop-up casino bonus offer was evaluated by two responsible gambling officials, four academics, and one ex-problem gambler. The use of the pop-up online casino bonus served as a useful template to evaluate the effectiveness of *explainable persuasion* in building resistance against persuasive design techniques used in online gambling platforms.

Another consideration is related to external validity. The findings are based on a controlled experiment, and in real life, individuals may respond differently, or not at all, to *explainable persuasion*. Due to the immersion effect of gambling, users may overlook explanations in real life and lose the ability to perceive external stimuli (Schüll, 2012; Murch et al., 2020).

6. Conclusion and future research

The findings of this study suggest that *explainable persuasion* may increase awareness of the presence and risks of

persuasive interfaces and strengthen user resistance if engaging with such persuasive interfaces is not aligned with personal goals. Moreover, the findings showed that *explainable persuasion* has the potential to function as both a preventative and a corrective approach for protecting users in the online gambling domain. The findings of the inoculation study showed that *explainable persuasion* could be a cost-effective way to sustain resilience against persuasive interfaces and attenuate excessive digital behavior.

We want to emphasize that we do not assert the ease of implementing *explainable persuasion*, considering the challenge of striking a balance between benefits, business objectives, and user experience. Our study aimed to take an initial step in evaluating the feasibility of using *explainable persuasion* as an intervention to encourage more responsible digital behavior. Future research could examine how to deliver explanations based on factors related to attention switching, maintenance, and communication processing. For example, research has shown that messages designed to encourage players to reflect, self-evaluate, and self-regulate are more effective than those that focus on informing players of the hazards associated with gambling. Future research also needs to examine the inoculation effect of *explainable persuasion* in other domains that utilize persuasive interfaces, such as social media or online streaming platforms. Furthermore, additional research is necessary to explore the design considerations unique to each persuasion principle and to grasp their individual effects.

Disclosure statement

No potential conflict of interest was reported by the author(s).

ORCID

Deniz Cemiloglu http://orcid.org/0000-0003-4725-3868
Raian Ali http://orcid.org/0000-0002-5285-7829

References

Albayram, Y., Khan, M. M. H., & Fagan, M. (2017). A study on designing video tutorials for promoting security features: A case study in the context of two-factor authentication (2FA). *International Journal of Human–Computer Interaction*, *33*(11), 927–942. https://doi.org/10.1080/10447318.2017.1306765

Ali, R., Jiang, N., Phalp, K., Muir, S., & McAlaney, J. (2015). *The emerging requirement for digital addiction labels*. Cham.

Alrobai, A., Phalp, K., & Ali, R. (2014). *Digital addiction: A requirements engineering perspective*. Requirements Engineering: Foundation for Software Quality.

Amazeen, M. A. (2021). Resisting covert persuasion in digital news: Comparing inoculation and reactance in the processing of native advertising disclosures and in article engagement intentions. *Journalism & Mass Communication Quarterly*, *98*(4), 1129–1156. https://doi.org/10.1177/1077699020952131

Amazeen, M. A., & Vargo, C. J. (2021). Sharing native advertising on Twitter: Content analyses examining disclosure practices and their inoculating influence. *Journalism Studies*, *22*(7), 916–933. https://doi.org/10.1080/1461670X.2021.1906298

Amazeen, M. A., & Wojdynski, B. W. (2019). Reducing native advertising deception: Revisiting the antecedents and consequences of persuasion knowledge in digital news contexts. *Mass Communication and Society*, *22*(2), 222–247. https://doi.org/10.1080/15205436.2018.1530792

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders: DSM-5*. APA.

Atkinson, B. M. C. (2006). *Captology: A critical review*. Persuasive Technology.

Auer, M. M., & Griffiths, M. D. (2016). Personalized behavioral feedback for online gamblers: A real world empirical study. *Frontiers in Psychology*, *7*, 1875. https://doi.org/10.3389/fpsyg.2016.01875

Banas, J. A., & Miller, G. (2013). Inducing resistance to conspiracy theory propaganda: Testing inoculation and metainoculation strategies. *Human Communication Research*, *39*(2), 184–207. https://doi.org/10.1111/hcre.12000

Banas, J. A., & Rains, S. A. (2010). A meta-analysis of research on inoculation theory. *Communication Monographs*, *77*(3), 281–311. https://doi.org/10.1080/03637751003758193

Barral, O., Aranyi, G., Kouider, S., Lindsay, A., Prins, H., Ahmed, I., Jacucci, G., Negri, P., Gamberini, L., Pizzi, D., & Cavazza, M. (2014). *Covert persuasive technologies: Bringing subliminal cues to human-computer interaction*. Persuasive Technology.

Berdichevsky, D., & Neuenschwander, E. (1999). Toward an ethics of persuasive technology. *Communications of the ACM*, *42*(5), 51–58. https://doi.org/10.1145/301353.301410

Bjørseth, B., Simensen, J. O., Bjørnethun, A., Griffiths, M. D., Erevik, E. K., Leino, T., & Pallesen, S. (2020). The effects of responsible gambling pop-up messages on gambling behaviors and cognitions: A systematic review and meta-analysis. *Frontiers in Psychiatry*, *11*, 601800. https://doi.org/10.3389/fpsyt.2020.601800

Blaszczynski, A., Collins, P., Fong, D., Ladouceur, R., Nower, L., Shaffer, H. J., Tavares, H., & Venisse, J.-L. (2011). Responsible gambling: General principles and minimal requirements. *Journal of Gambling Studies*, *27*(4), 565–573. https://doi.org/10.1007/s10899-010-9214-0

Blaszczynski, A., Ladouceur, R., Nower, L., & Shaffer, H. (2013). Informed choice and gambling: Principles for consumer protection. *The Journal of Gambling Business and Economics*, *2*(1), 103–118. https://doi.org/10.5750/jgbe.v2i1.527

Blaszczynski, A., Ladouceur, R., & Shaffer, H. J. (2004). A science-based framework for responsible gambling: The Reno model. *Journal of Gambling Studies*, *20*(3), 301–317. https://doi.org/10.1023/B:JOGS.0000040281.49444.e2

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Cacioppo, J. T., & Petty, R. E. (1981). Social psychological procedures for cognitive response assessment: The thought-listing technique. *Cognitive Assessment*, 309–342.

Cacioppo, J. T., Von Hippel, W., & Ernst, J. M. (1997). Mapping cognitive structures and processes through verbal content: The thought-listing technique. *Journal of Consulting and Clinical Psychology*, *65*(6), 928–940. https://doi.org/10.1037/0022-006X.65.6.928

Caillon, J., Grall-Bronnec, M., Saillard, A., Leboucher, J., Péré, M., & Challet-Bouju, G. (2021). Impact of warning pop-up messages on the gambling behavior, craving, and Cognitions of online gamblers: A randomized controlled trial. *Frontiers in Psychiatry*, *12*, 711431. https://doi.org/10.3389/fpsyt.2021.711431

Cameron, K. A. (2009). A practitioner's guide to persuasion: An overview of 15 selected persuasion theories, models and frameworks. *Patient Education and Counseling*, *74*(3), 309–317. https://doi.org/10.1016/j.pec.2008.12.003

Cantinotti, M., Ladouceur, R., & Jacques, C. (2004). Sports betting: Can gamblers beat randomness? *Psychology of Addictive Behaviors*, *18*(2), 143–147. https://doi.org/10.1037/0893-164X.18.2.143

Cemiloglu, D., Arden-Close, E., Hodge, S. E., & Ali, R. (2023). Explainable persuasion for interactive design: The case of online gambling. *Journal of Systems and Software*, *195*, 111517. https://doi.org/10.1016/j.jss.2022.111517

Cemiloglu, D., Catania, M., & Ali, R. (2021). Explainable persuasion in interactive design. In *2021 IEEE 29th International Requirements*

Engineering Conference Workshops (REW), https://doi.org/10.1109/REW53955.2021.00066
Cemiloglu, D., Naiseh, M., Catania, M., Oinas-Kukkonen, H., & Ali, R. (2021). *The fine line between persuasion and digital addiction*. Persuasive Technology.
Compton, J. (2013). Inoculation theory. In *The SAGE handbook of persuasion: Developments in theory and practice* (Vol. 2, pp. 220–237). SAGE.
Compton, J. (2020). Prophylactic versus therapeutic inoculation treatments for resistance to influence. *Communication Theory*, *30*(3), 330–343. https://doi.org/10.1093/ct/qtz004
Compton, J., & Ivanov, B. (2012). Untangling threat during inoculation-conferred resistance to influence. *Communication Reports*, *25*(1), 1–13. https://doi.org/10.1080/08934215.2012.661018
Compton, J., & Pfau, M. (2004). Use of inoculation to foster resistance to credit card marketing targeting college students. *Journal of Applied Communication Research*, *32*(4), 343–364. https://doi.org/10.1080/0090988042000276014
Compton, J., & Pfau, M. (2008). Inoculating against pro-plagiarism justifications: Rational and affective strategies. *Journal of Applied Communication Research*, *36*(1), 98–119. https://doi.org/10.1080/00909880701799329
Compton, J., & Pfau, M. (2009). Spreading inoculation: Inoculation, resistance to influence, and word-of-mouth communication. *Communication Theory*, *19*(1), 9–28. https://doi.org/10.1111/j.1468-2885.2008.01330.x
Davis, J. (2009). Design methods for ethical persuasive computing. In *Proceedings of the 4th International Conference on Persuasive Technology*. https://doi.org/10.1145/1541948.1541957
de Pelsmacker, P., & Neijens, P. C. (2012). New advertising formats: How persuasion knowledge affects consumer responses. *Journal of Marketing Communications*, *18*(1), 1–4. https://doi.org/10.1080/13527266.2011.620762
Devlin, M. E., & Walton, D. (2012). The prevalence of problem gambling in New Zealand as measured by the PGSI: Adjusting prevalence estimates using meta-analysis. *International Gambling Studies*, *12*(2), 177–197. https://doi.org/10.1080/14459795.2011.653384
Encyclopædia Britannica. (1998). *Slot machine*. Retrieved May 15, from https://www.britannica.com/topic/slot-machine.
Faden, R. R., & Beauchamp, T. L. (1986). *A history and theory of informed consent*. Oxford University Press.
Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146
Ferris, J. A., & Wynne, H. J. (2001a). *The Canadian problem gambling index*. Canadian Centre on Substance Abuse.
Ferris, J. A., & Wynne, H. J. (2001b). *The Canadian problem gambling index: User manual*. Canadian Centre on Substance Abuse.
Fogg, B. J. (2003). *Persuasive technology: using computers to change what we think and do*. Morgan Kaufmann. https://doi.org/10.1145/764008.763957
Fransen, M. L., Smit, E. G., & Verlegh, P. W. (2015). Strategies and motives for resistance to persuasion: An integrative framework. *Frontiers in Psychology*, *6*(1), 1201. https://doi.org/10.3389/fpsyg.2015.01201
Fransen, M. L., Verlegh, P. W., Kirmani, A., & Smit, E. G. (2015). A typology of consumer strategies for resisting advertising, and a review of mechanisms for countering them. *International Journal of Advertising*, *34*(1), 6–16. https://doi.org/10.1080/02650487.2014.995284
Friestad, M., & Wright, P. (1994). The Persuasion knowledge model: How people cope with persuasion attempts. *Journal of Consumer Research*, *21*(1), 1–31. https://doi.org/10.1086/209380
Gidron, Y., Levy, E., Farchi, M., & Rapaport, C. (2023). Effects of an automatized psychological inoculation (PI) intervention on anxiety, resilience and adherence to COVID-19 recommendations. *Psychology & Health*, *38*(5), 541–554. https://doi.org/10.1080/08870446.2021.1984485
Godbold, L. C., & Pfau, M. (2000). Conferring resistance to peer pressure among adolescents: Using inoculation theory to discourage alcohol use. *Communication Research*, *27*(4), 411–437. https://doi.org/10.1177/009365000027004001
Goldhaber, M. H. (1997). The attention economy and the net. *First Monday*, *2*(4), 519. https://doi.org/10.5210/fm.v2i4.519
Graml, T., Loock, C.-M., Baeriswyl, M., & Staake, T. (2011). Improving residential energy consumption at large using persuasive systems. *19th European Conference on Information Systems, ECIS 2011*.
Hinson, J. M., Whitney, P., Holben, H., & Wirick, A. K. (2006). Affective biasing of choices in gambling task decision making. *Cognitive, Affective & Behavioral Neuroscience*, *6*(3), 190–200. https://doi.org/10.3758/CABN.6.3.190
Hogan, E. A. (2001). The attention economy: Understanding the New Currency of Business. *Academy of Management Perspectives*, *15*(4), 145–147. https://doi.org/10.5465/ame.2001.5898765
Insko, C. A. (1967). *Theories of attitude change*. Appleton.
Ivanov, B., Parker, K. A., & Dillingham, L. L. (2018). Testing the limits of inoculation-generated resistance. *Western Journal of Communication*, *82*(5), 648–665. https://doi.org/10.1080/10570314.2018.1454600
Ivanov, B., Pfau, M., & Parker, K. A. (2009). The attitude base as a moderator of the effectiveness of inoculation strategy. *Communication Monographs*, *76*(1), 47–72. https://doi.org/10.1080/03637750802682471
Jones, J. G., & Simons, H. W. (2017). *Persuasion in society*. Routledge.
Kahneman, D. (2011). *Thinking, fast and slow*. Macmillan.
Kiesler, C. A., & Kiesler, S. B. (1964). Role of forewarning in persuasive communications. *Journal of Abnormal Psychology*, *68*(5), 547–549. https://doi.org/10.1037/h0042145
Kim, S., Wogalter, M. S. (2009). Habituation, dishabituation, and recovery effects in visual warnings. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*,
Knowles, E. S., & Linn, J. A. (2004). *Resistance and persuasion*. Taylor & Francis Inc.
Korn, D. A., & Shaffer, H. J. (1999). Gambling and the health of the public: Adopting a public health perspective. *Journal of Gambling Studies*, *15*(4), 289–365. https://doi.org/10.1023/A:1023005115932
Kuonanoja, L., & Oinas-Kukkonen, H. (2018). *Recognizing and mitigating the negative effects of information technology use: A systematic review of persuasive characteristics in information systems*. Nordic Contributions in IS Research. https://doi.org/10.1145/3290605.3300616
Langer, E. J. (1975). The illusion of control. *Journal of Personality and Social Psychology*, *32*(2), 311–328. https://doi.org/10.1037/0022-3514.32.2.311
Langrial, S., Lehto, T., Oinas-Kukkonen, H., Harjumaa, M., & Karppinen, P. (2012). *Native mobile applications for personal well-being: a persuasive systems design evaluation* PACIS 2012 Proceeding.
Levy, E., Gidron, Y., Deschepper, R., Olley, B. O., & Ponnet, K. (2019). Effects of a computerized psychological inoculation intervention on condom use tendencies in sub Saharan and Caucasian students: Two feasibility trials. *Health Psychology and Behavioral Medicine*, *7*(1), 160–178. https://doi.org/10.1080/21642850.2019.1614928
Livingstone, S., & Helsper, E. J. (2006). Does advertising literacy mediate the effects of advertising on children? A critical examination of two linked research literatures in relation to obesity and food choice. *Journal of Communication*, *56*(3), 560–584. https://doi.org/10.1111/j.1460-2466.2006.00301.x
Lopez-Gonzalez, H., Estévez, A., & Griffiths, M. D. (2018). Spanish validation of the Problem Gambling Severity Index: A confirmatory factor analysis with sports bettors. *Journal of Behavioral Addictions*, *7*(3), 814–820. https://doi.org/10.1556/2006.7.2018.84
Lumsdaine, A. A., & Janis, I. L. (1953). Resistance to "counterpropaganda" produced by one-sided and two-sided "propaganda" presentations. *Public Opinion Quarterly*, *17*(3), 311–318. https://doi.org/10.1086/266464
Mason, A. M., & Miller, C. H. (2013). Inoculation message treatments for curbing noncommunicable disease development. *Revista*

panamericana de salud publica = Pan American Journal of Public Health, *34*(1), 29–35.
McCormack, A., & Griffiths, M. D. (2013). A scoping study of the structural and situational characteristics of internet gambling. *International Journal of Cyber Behavior, Psychology and Learning*, *3*(1), 29–49. https://doi.org/10.4018/ijcbpl.2013010104
McGuier, W. (1962). Persistence of the resistance to persuasion induced by various types of prior belief defenses. *The Journal of Abnormal and Social Psychology*, *64*(4), 241–248. https://doi.org/10.1037/h0044167
McGuire, W. J. (1961). Resistance to persuasion conferred by active and passive prior refutation of the same and alternative counterarguments. *The Journal of Abnormal and Social Psychology*, *63*(2), 326–332. https://doi.org/10.1037/h0048344
McGuire, W. J. (1964). Inducing resistance to persuasion. Some contemporary approaches. In C. C. Haaland & W. O. Kaelber (Eds.), *Self and society. An anthology of readings* (pp. 192–230). Ginn Custom Publishing 1981.
McGuire, W. J., & Papageorgis, D. (1962). Effectiveness of forewarning in developing resistance to persuasion. *Public Opinion Quarterly*, *26*(1), 24–34. https://doi.org/10.1086/267068
Monaghan, S. (2009). Responsible gambling strategies for Internet gambling: The theoretical and empirical base of using pop-up messages to encourage self-awareness. *Computers in Human Behavior*, *25*(1), 202–207. https://doi.org/10.1016/j.chb.2008.08.008
Monaghan, S., & Blaszczynski, A. (2010). Electronic gaming machine warning messages: Information versus self-evaluation. *The Journal of Psychology*, *144*(1), 83–96. https://doi.org/10.1080/00223980903356081
Murch, W. S., Limbrick-Oldfield, E. H., Ferrari, M. A., MacDonald, K. I., Fooken, J., Cherkasova, M. V., Spering, M., & Clark, L. (2020). Zoned in or zoned out? Investigating immersion in slot machine gambling using mobile eye-tracking. *Addiction, 115*(6), 1127–1138. https://doi.org/10.1111/add.14899
Oinas-Kukkonen, H. (2013). A foundation for the study of behavior change support systems. *Personal and Ubiquitous Computing*, *17*(6), 1223–1235. https://doi.org/10.1007/s00779-012-0591-5
Oyebode, O., Ndulue, C., Alhasani, M., & Orji, R. (2020). *Persuasive mobile apps for health and wellness: A comparative systematic review*. International Conference on Persuasive Technology.
Panic, K., Cauberghe, V., & De Pelsmacker, P. (2013). Comparing TV ads and advergames targeting children: The impact of persuasion knowledge on behavioral responses. *Journal of Advertising*, *42*(2–3), 264–273. https://doi.org/10.1080/00913367.2013.774605
Parker, K. A., Ivanov, B., & Compton, J. (2012). Inoculation's efficacy with young adults' risky behaviors: Can inoculation confer cross-protection over related but untreated issues? *Health Communication*, *27*(3), 223–233. https://doi.org/10.1080/10410236.2011.575541
Petty, R. E., & Cacioppo, J. T. (1979). Effects of forwarning of persuasive intent and involvement on cognitive responses and persuasion. *Personality and Social Psychology Bulletin*, *5*(2), 173–176. https://doi.org/10.1177/014616727900500209
Petty, R. E., & Cacioppo, J. T. (2012). *Communication and persuasion: Central and peripheral routes to attitude change*. Springer Science & Business Media.
Pfau, M., Bockern, S. V., & Kang, J. G. (1992). Use of inoculation to promote resistance to smoking initiation among adolescents. *Communication Monographs*, *59*(3), 213–230. https://doi.org/10.1080/03637759209376266
Pfau, M., & Burgoon, M. (1988). Inoculation in political campaign communication. *Human Communication Research*, *15*(1), 91–111. https://doi.org/10.1111/j.1468-2958.1988.tb00172.x
Pfau, M., Compton, J., Parker, K., Wittenberg, E., An, C., Ferguson, M., & Malyshev, Y. (2004). The traditional explanation for resistance based on the core elements of threat and counterarguing and an alternative rationale based on attitude accessibility: Do these mechanisms trigger distinct or overlapping process of resistance. *Human Communication Research*, *30*(3), 329–360. https://doi.org/10.1093/hcr/30.3.329
Pfau, M., Haigh, M., Fifrick, A., Holl, D., Tedesco, A., Cope, J., Nunnally, D., Schiess, A., Preston, D., Roszkowski, P., & Martin, M. (2006). The effects of print news photographs of the casualties of war. *Journalism & Mass Communication Quarterly*, *83*(1), 150–168. https://doi.org/10.1177/107769900608300110
Pfau, M., Park, D., Holbert, R. L., & Cho, J. (2001). The effects of party-and PAC-sponsored issue advertising and the potential of inoculation to combat its impact on the democratic process. *American Behavioral Scientist*, *44*(12), 2379–2397. https://doi.org/10.1177/00027640121958384
Pfau, M., Semmler, S. M., Deatrick, L., Mason, A., Nisbett, G., Lane, L., Craig, E., UNDerhill, J., & Banas, J. (2009). Nuances about the role and impact of affect in inoculation. *Communication Monographs*, *76*(1), 73–98. https://doi.org/10.1080/03637750802378807
Pfau, M., Szabo, A., Anderson, J., Morrill, J., Zubric, J., & H-Wan, H. H. (2001). The role and impact of affect in the process of resistance to persuasion. *Human Communication Research*, *27*(2), 216–252. https://doi.org/10.1111/j.1468-2958.2001.tb00781.x
Pfau, M., Tusing, K. J., Koerner, A. F., Lee, W., Godbold, L. C., Penaloza, L. J., Yang, V. S.-H., & Hong, Y.-H. (1997). Enriching the inoculation construct: The role of critical components in the process of resistance. *Human Communication Research*, *24*(2), 187–215. https://doi.org/10.1111/j.1468-2958.1997.tb00413.x
Reader's Digest. (2020). *The History of Gambling*. Retrieved May, 2022, from https://www.readersdigest.co.uk/inspire/life/the-history-of-gambling.
Reynolds-Tylus, T., Martinez Gonzalez, A., & Quick, B. L. (2019). The role of choice clustering and descriptive norms in attenuating psychological reactance to water and energy conservation messages. *Environmental Communication*, *13*(7), 847–863. https://doi.org/10.1080/17524032.2018.1461672
Roloff, M. E., & Miller, G. R. (1980). *Persuasion: New directions in theory and research (Vol. 8)*. SAGE Publications, Incorporated.
Roozenbeek, J., & van der Linden, S. (2019a). Fake news game confers psychological resistance against online misinformation. *Palgrave Communications*, *5*(1), 1–10. https://doi.org/10.1057/s41599-019-0279-9
Roozenbeek, J., & Van Der Linden, S. (2019b). The fake news game: Actively inoculating against the risk of misinformation. *Journal of Risk Research*, *22*(5), 570–580. https://doi.org/10.1080/13669877.2018.1443491
Schüll, N. D. (2012). *Addiction by design*. Princeton University Press.
Sheskin, D. J. (2003). *Handbook of parametric and nonparametric statistical procedures*. Chapman and Hall/CRC.
Smids, J. (2012). *The Voluntariness of Persuasive Technology Persuasive Technology*. Design for Health and Safety.
So, R., Matsushita, S., Kishimoto, S., & Furukawa, T. A. (2019). Development and validation of the Japanese version of the problem gambling severity index. *Addictive Behaviors*, *98*, 105987. https://doi.org/10.1016/j.addbeh.2019.05.011
Spagnolli, A., Chittaro, L., & Gamberini, L. (2016). Interactive persuasive systems: A perspective on theory and evaluation. *International Journal of Human-Computer Interaction*, *32*(3), 177–189. https://doi.org/10.1080/10447318.2016.1142798
Spahn, A. (2012). And lead us (not) into persuasion…? Persuasive technology and the ethics of communication. *Science and Engineering Ethics*, *18*(4), 633–650. https://doi.org/10.1007/s11948-011-9278-y
Timmer, J., Kool, L., & van Est, R. (2015). *Ethical challenges in emerging applications of persuasive technology*. Persuasive Technology.
van der Linden, S., Roozenbeek, J., Greifeneder, R., Jaffe, M., Newman, E., & Schwarz, N. (2020). Psychological inoculation against fake news. *The Psychology of Fake News*, 147–169. https://doi.org/10.4324/9780429295379-11
Walker, M. B. (1992). Irrational thinking among slot machine players. *Journal of Gambling Studies*, *8*(3), 245–261. https://doi.org/10.1007/BF01014652
Wojdynski, B. W., & Golan, G. J. (2016). Native advertising and the future of mass communication. *American Behavioral Scientist*, *60*(12), 1403–1407. https://doi.org/10.1177/0002764216660134
Wright, P. (1975). Factors affecting cognitive resistance to advertising. *Journal of Consumer Research*, *2*(1), 1–9. https://doi.org/10.1086/208610

Zaichkowsky, J. L. (1985). Measuring the involvement construct. *Journal of Consumer Research*, *12*(3), 341–352. https://doi.org/10.1086/208520

Zhang-Kennedy, L., & Chiasson, S. (2022). A systematic review of multimedia tools for cybersecurity awareness and education. *ACM Computing Surveys*, *54*(1), 1–39. https://doi.org/10.1145/3427920

Zuwerink Jacks, J., & Cameron, K. A. (2003). Strategies for resisting persuasion. *Basic and Applied Social Psychology*, *25*(2), 145–161. https://doi.org/10.1207/S15324834BASP2502_5

About the authors

Deniz Cemiloglu completed her PhD at the Department of Computing and Informatics, Bournemouth University, Poole, UK. Her current research areas are persuasive technology, explainability, digital wellness and responsibility by design. Her work focuses on how to elicit, document, adapt, prioritize, and validate ethical requirements to embed responsibility in potentially addictive technology.

Selin Gurgun is a PhD candidate in the Computing and Informatics department at Bournemouth University, specializing in the field of social media design and its impact on challenging misinformation. Her research focuses on understanding the dynamics of interpersonal relationships in the online environment, with a particular interest in Human-Computer Interaction and social corrections.

Emily Arden-Close is a Principal Academic in Psychology in the Department of Psychology, Bournemouth University, Poole, UK, and a registered Health Psychologist with the Health & Care Professions Council, UK. Her research focuses on the design, development and evaluation of online behavior change interventions to maintain well-being in individuals with long-term health conditions.

Nan Jiang is the acting Head of Computing and Informatics and Professor in Human Computer Interaction. He joined BU in 2010. Prior to this, Nan was an Open Distance Learning (ODL) tutor at Queen Mary, University of London, where he also received his PhD in Web Usability and MSc in Advanced Methods.

Raian Ali is a Professor at the College of Science and Engineering, Hamad Bin Khalifa University, Qatar. His research focuses on the design process of digital motivation solutions, such as Gamification and Persuasive Technology, to aid people change or maintain their attitude and behavior in areas like health and enterprise productivity.