The Misinformation Game: The Effectiveness of Analogue Serious Games as an Inoculation Against Misinformation

Niklas Henderson

Submitted for the degree of Doctor of Philosophy

University of East Anglia, School of Computing Sciences

Submitted September 2025

Abstract

Misinformation poses a serious risk to societies, and its rapid spread particularly through online platforms has negatively impacted democratic processes, public health campaigns, and more. Game-based inoculation theory has been proposed as an engaging individual-level communication strategy to improve misinformation resilience in individuals, but much remains unknown of the approach's effectiveness over time and the underlying mechanisms conferring resistance in this context. In response to these research gaps, this thesis uses the collaborative card game *Fake News* (Roozenbeek & van der Linden, 2018) to explore how players learn about misinformation from game-based interventions and how intervention design features shape these outcomes. This thesis also more deeply interrogates inoculation theory itself, assessing the theory's present application in game-based contexts and exploring how the theory's mechanisms may be most effectively elicited through serious games, culminating in the development of a new collaborative card game against digital social engineering, *The disPHISHinformation Game*. As well as a within-subjects study design using quantitative measures to assess intervention effectiveness (N = 118), this thesis takes the novel step of integrating qualitative methods, namely participant observation and semi-structured interviews, to more effectively explore how players experience game-based interventions both individually and collaboratively.

Quantitative findings from the first study showed no significant improvements in participants' resilience to 'fake news' after playing *Fake News*, a result that contrasts with previous findings using the same intervention. Conversely, qualitative insights showed how the intervention's tangible, embodied, and analogue design provided a platform for collaborative and experiential learning as players completed the intervention together. These findings indicate the need for a wider consideration of learning outcomes from game-based interventions and highlight the value of mixed-method research design in this context. An integrative narrative review of inoculation theory's application in game-based contexts also supported auxiliary findings from research with *Fake News*. The review shows that despite being a core foundational theory of many game-based interventions, how inoculation theory has informed intervention design, and in particular how inoculation's mechanisms are elicited, is often superficially defined. Informed by these findings, this thesis culminates in the

design and testing of the new collaborative card game *The disPHISHinformation Game*, showcasing a design approach that more carefully considers inoculation theory's role in the game design process. This thesis demonstrates the importance and value of mixed-method approaches, gives insights into the broad effects of these game-based interventions, and explores both the limitations of inoculation theory's current interpretation in game-based interventions, and how it may be more fully incorporated in this context.

Contents

List of Publications

Henderson, N., Pallett, H., van der Linden, S., Roozenbeek, J. (2025) **Exploring the Learning Potential of a Game-based Intervention to Inoculate Against Misinformation.** *Social Media + Society*. [In Preparation]

Henderson, N., & Pallett, H. (2026). **Inoculation theory as a design approach to game-based misinformation interventions: a review**. *Popular Communication*, 1–23. https://doi.org/10.1080/15405702.2026.2619473

Henderson, N., Pallett, H., van der Linden, S., Montanarini, J., & Buckley, O. (2024). **The disPHISHinformation Game: Creating a Serious Game to Fight Phishing Using Blended Design Approaches**. *Human Factors in Cybersecurity*, *127*, 146–156. https://doi.org/10.54941/ahfe1004774

Henderson, N., Buckley, O., & Pallett, H. (2023). **Investigating Longitudinal Effects of Physical Inoculation Interventions Against Disinformation**. In C. Stephanidis, M. Antona, S. Ntoa, & G. Salvendy (Eds.), *HCI International 2023 Posters* (pp. 39–46). Springer Nature Switzerland. https://doi.org/https://doi.org/10.1007/978-3-031-36001-*5_5*

List of Tables

List of Figures

Acknowledgements

I would first like to acknowledge my primary supervisor Helen Pallett, who has introduced me to a whole new world of research with patience and encouragement. Your constant support and constructive feedback have been invaluable throughout this project; my PhD journey has not at all been what I had expected, and I am hugely grateful for the impactful influence you have had on it.

I would also like to extend my gratitude to everyone who took part in this research, including participants in both rounds of data collection, and especially those who shared their time and insights with me during interviews. I would also like to thank AVIVA for their time and openness in being a part of this research, and those at AVIVA who took part in playtesting the intervention. To this end, I would like to thank my friends and family for their input and help with not only the cyber security serious game, but for their support throughout this thesis.

I was also very fortunate to be able to spend 6 months as a visiting scholar at the Social Decision-Making Lab at the University of Cambridge with Professor Sander van der Linden and Dr Jon Roozenbeek. This time was a hugely transformative experience for me in not only learning but inspiration, and I would like to extend my deepest thanks to Sander, Jon, and all the other members I got to know in the lab during my stay.

1. Introduction

In 2024, the United Kingdom experienced riots and 'anti-immigration protests' in the aftermath of the murder of three young girls in Southport on the 29th of July. News was shared online that a Muslim asylum-seeker 'Ali Al Shakati' (Burnell & PA Media, 2024) was responsible for the killing, and the riots, fuelled by posts online by far-right groups and politicians, led to the destruction and looting of property and shops, the injuring of over 350 police officers, and the arrest of over 1,500 individuals. The speed and intensity with which the riots took place shocked many in the UK and led to many counterdemonstrations as communities tried to show their support for those affected by the riots (Bonnett & Hopkins, 2025). Despite the police issuing a statement that this shared information was not correct, many users continued to interact with misinformation online (Cheshire & Doak, 2024). A public inquiry launched by the UK government highlighted how information on the killer, particularly his ethnicity and religion, had spread through both social media platforms such as Facebook and X (formerly Twitter), and on closed groups on encrypted platforms such as Telegram and WhatsApp. Only, this information was not true; this misinformation of the name, ethnicity, and religion of the perpetrator was posted by the website Channel3Now, a news channel that "resembles Russian approaches around information laundering and narrative dissemination" (Logically, 2024). This example represents one of many global incidents that have been tied to misinformation (with often tragic consequences, see BBC, 2020; M. Fisher et al., 2016; Jolley & Paterson, 2020).

Research indicates that more than 9 in 10 people in the UK have reported reading misinformation online (Enock et al., 2024), and a recent report by Ofcom has demonstrated that UK adults are concerned about misinformation and its societal impact (Ofcom, 2024b; Altay, 2023; Newman et al., 2024; Office for National Statistics, 2014). The risks that misinformation pose to not only safety but wider issues such as geopolitics and democracy has never been greater (Lewandowsky et al., 2023). Democracy relies on a well-informed populus (Kuklinski et al., 2000), but it is at risk. Scholars have highlighted democratic backsliding (i.e., moving away from democratic ideals) is increasing globally (Abels et al., 2024; Lewandowsky et al., 2025).

Misinformation is not a new phenomenon. History gives many examples of how misinformation shaped the societies in which we live today, from the defamatory will and testament of Mark Antony in ancient Rome (historians have long discussed

whether Octavian may have forged the document, see J. R. Johnson, 1978; Sirianni, 1984) to exaggerated and false claims around the serial killer Jack the Ripper in Victorian London (Morrison, 2019). Some have argued that misinformation has always played a role in the information environment of human communication and thus claims of 'post-truth' (Kucharski, 2016) or crisis (Lewandowsky et al., 2023) are overstated (Altay et al., 2023; C. W. Anderson, 2021; Carlson, 2020; Roozenbeek & van der Linden, 2024). This may be misleading, however; the (digital) landscape in which individuals consume news means that misinformation moves faster and is more impactful than ever before (Roozenbeek & van der Linden, 2024; van der Linden & Kyrychenko, 2024).

The characteristics of content and modes of distribution for news has changed drastically through the 'boom' of the internet (Roozenbeek & van der Linden, 2024). The accessibility and subsequent popularity of the world wide web have drastically changed the incentives for creators of news. Advertising is by far the dominant revenue stream for many popular social media platforms (Meta Platforms Inc, 2024), and the cross-border reach of social media has allowed not only individuals but malicious states to generate profit (Center for Countering Digital Hate [CCDH], 2024; McLoughlin et al., 2024; Rathje & van der Linden, 2023) and influence public opinion and democratic processes (Bessi & Ferrara, 2016; Erizanu, 2024) with little concern for regional legal jurisdiction. Social media platforms have also facilitated the rise of citizen journalism which, although giving global voice to individuals not previously possible, can also facilitate the spread of misinformation to large audiences at a faster rate than ever before. Social media platforms themselves have little incentive to moderate uploaded content (Silverman, 2025); many platforms such as Meta generate revenue primarily through advertising, meaning that business incentives have moved away from quality and instead focus on attention (J. Williams, 2018). Research has shown that expression of morality and emotion is the greatest predictor of engagement, and thus user traffic (Brady et al., 2017, 2020; Valenzuela et al., 2017).

This issue is exacerbated by users: humans are not unbiased information processing machines. Research has shown that users view information through the lens of many biases, drawing on users' prior experiences and held beliefs. Two leading theories have been proposed: classical attention, or motivated cognition. The classical attention account of misinformation belief posits that the leading reason users fall for a false belief is a lack of motivation (or 'laziness', see Pennycook & Rand, 2019) to evaluate read information. The context of social media and volume of content has been demonstrated to influence users into using intuitive cognitive mechanisms, rather than deliberate decision making on information veracity (Bago et al., 2020; Bronstein et al., 2019; Evans, 2003; Swami et al., 2014). The second account of misinformation belief is one of motivated cognition (Kunda, 1990). Given a user's existing beliefs, affiliations,

and motivations, they evaluate information with the intention to protect this identity. The types of commitments that people have to their affinity groups (e.g., political, religious, social identity) has been shown to lead to selective media endorsement (Bolsen et al., 2014; Kahan et al., 2017; Osmundsen et al., 2021; van der Linden et al., 2021). For example, an individual with strong affiliation to a political party may be motivated to evaluate information veracity depending on whether or not it aligns with the beliefs or public stance of that political group (Osmundsen et al., 2021). Some scholars believe that both inattention and motivated cognition contribute to misinformation belief (see 'integrative' model of misinformation belief, Van Bavel et al., 2021).

Research has demonstrated, however, that users' ability to correctly discern the veracity of information can be improved. In particular, individual-level interventions have been shown to improve resilience to misinformation. Educational approaches such as media literacy training and promoting critical thinking have broad literature demonstrating their effectiveness (Axelsson et al., 2021; Guess et al., 2020a; Nygren & Guath, 2019). Debunking (i.e., correcting mistruths after they are read or believed) is a common approach which meta-analyses have shown to be effective (Martel & Rand, 2024; Walter et al., 2020), although fact checking on social media platforms requires a willingness to engage from the platform themselves (McMahon et al., 2025; Silverman, 2025). More effective, however, is pre-bunking, the act of *pre-emptively* refuting misinformation. Pre-emptively warning an individual of mistruths they may encounter in the future helps build the individual's resilience to it, a process which is often easier than debunking once individual is already misinformed. Inoculation theory (sometimes referred to as "the grandfather theory to persuasion", Eagly & Chaiken, 1993, p. 561), has been proposed as a leading communication approach to pre-emptively building resistance to misinformation in individuals, and has attracted much scholarly interest particularly in the last decade (Compton, 2024).

Inoculation theory posits that, on a given issue (on which an individual's belief will be attacked), by a) making the individual aware that their correct belief is vulnerable to attack, and b) providing them with counterarguments on the issue, they will be better-equipped, and more importantly better-motivated, to defend this belief from future persuasive attack. A meta-analysis (Banas & Rains, 2010) has demonstrated inoculation theory to be an effective method to building resilience to misinformation (as well as achieving other outcomes such as reducing sharing intentions, see Lu et al., 2023), and the mechanisms, components, and moderators are supported by a large body of body of literature (Compton, 2024). Inoculation hesitancy (i.e., the reluctance of individuals to engage with inoculation interventions) remains a barrier for inoculation theory; an intervention may be efficacious, but low uptake will limit its

effectiveness (A. Johnson & Madsen, 2024; Roozenbeek et al., 2024). Recently, game-based inoculation interventions have been explored as, amongst other things, an approach to scale up inoculation and make it more attractive to engage with. Game-based interventions have been applied to a broad range of topics, such as misinformation on climate change, health, and politics, have won awards for innovation and communication, and have reached large audiences across the globe (for example, *Go Viral!* has been played over a million times since its launch in 2018, Lewsey, 2020).

However, the ability of game-based inoculation interventions to build resistance to misinformation in players has not yet been definitively established. Some interventions have produced small (Jeon et al., 2021) or null effect sizes (Harjani et al., 2023; Pomichal & Trnka, 2023; Rędzio et al., 2023) in preliminary research, and replication studies deploying an existing intervention in new (particularly global South) countries have also failed (Harjani et al., 2023). Moreover, very little is understood of the longitudinal effect of game-based interventions (Maertens et al., 2021), and of the little research on how conferred resistance decays over time, *learning* has emerged as a key variable (Maertens et al., 2025). Literature on learning has a long and rich history, however (Dewey, 1933, 1938; Mezirow, 1987; Schön, 1984), and this is yet to be comprehensively explored in the context of inoculation theory. The complex and interconnected nature of serious game design (i.e., the design of games with purposes beyond entertainment) has recently triggered calls for deeper research into why game-based intervention have been (in)effective (Appel et al., 2025; Harjani et al., 2023; Ma et al., 2023; Roozenbeek et al., 2024; Saleh et al., 2023). Moreover, there is emerging concern that serious games against misinformation under explore game design factors and mechanics (Kiili et al., 2024).

Addressing these concerns is essential for equipping scholars and practitioners with the tools necessary for creating more effective interventions and reducing the effect and spread of misinformation. This thesis aims to provide deeper and richer insights into the mechanics and effectiveness of game-based interventions against misinformation, in order to broaden the scope and provide new approaches to this rapidly expanding field. While existing studies in this literature tend to rely on quantitative and statistical methods alone to determine the effectiveness of game-based interventions, this thesis instead expands the methodological offer, including qualitative and engaged research methods and approaches which provide much-needed insights into how these interventions work, and how they could be made more effective. In this thesis I address these knowledge gaps through three central research questions around which the thesis is structured.

RQ1. How effective are analogue game-based interventions against misinformation, and to what extent does this change over time?

RQ2. What kinds of learning are stimulated by analogue game-based inoculation interventions against misinformation?

RQ3. How could game-based inoculation interventions against misinformation be better designed?

These research questions demand multi-disciplinary research which adopts novel theoretical and methodological approaches in order to provide rich new empirical insights. The studies in this thesis have followed an inductive pragmatist research philosophy with this aim, prioritising applicability and practicality over any one methodological approach. The first study, a mixed-methods within-subjects longitudinal study, explores the longitudinal efficacy and evidence of learning from the analogue intervention *Fake News*. Next, an integrative narrative review takes a cross-sectional assessment of current game-based inoculation interventions against misinformation, revealing design trends, practices, and processes. Finally, an exploratory study takes an engaged, considered approach to designing a new game-based intervention *The disPHISHinformation Game*, more comprehensively drawing on inoculation theory and (serious) game design literature. This thesis is structured as follows. First, the literature on interventions against misinformation is reviewed, contextualising game-based interventions in wider misinformation intervention literature, introducing inoculation theory and providing an overview of how it is operationalised in conventional and game-based contexts, and exploring links to theories of learning and game design, in order to explain and justify the research questions. Next, the research methods adopted in this thesis are described and justified. As well as discussing the research paradigms and design strategies adopted, this chapter gives an in-depth overview of the three studies presented in this thesis. Each study is introduced in relation to the research question, its methods comprehensively outlined and justified, and potential limitations discussed.

The first empirical chapter of this thesis (Chapter 4) outlines the quantitative results from a longitudinal within-subjects mixed-methods study with the analogue intervention *Fake News* (and represents a replication with extension of Roozenbeek & van der Linden, 2018). Focussing on the first research question, this chapter not only overviews the intervention's measured efficacy, but discusses socio-demographic moderators, the importance of participants' prior experiences with misinformation in relation to the intervention, and the leading role of intervention experience and design in measured outcomes.

The second empirical chapter of this thesis (Chapter 5) presents and discusses the qualitative findings from the study with *Fake News*. Following a pragmatist research paradigm, the study broke with convention by introducing qualitative elements to measure intervention efficacy, namely semi-structured interviews and session observations. With an emphasis on the second research question, this chapter studies participants' recollection, learning, and enjoyment from the intervention, further explores the impact of participant's prior experience with misinformation on the intervention, discusses qualitative evidence of inoculation mechanism and studies intervention outcomes through the lens of (serious) game design literature. This holistic evaluation of the intervention not only contributes to the literature but contextualises many quantitative findings from the first empirical chapter.

Following findings from the previous chapter on the role of inoculation theory in the analogue intervention *Fake News*, this chapter (Chapter 6) seeks to better understand how inoculation has been realised in game-based interventions, with the aim of contributing to the third research question. An integrative narrative review of 25 game-based inoculation interventions against misinformation is completed. The reviewing processes adopted includes not only the reviewing of publications, but importantly the playtesting of interventions by the researcher. The review reveals trends in not only intervention design, but how inoculation theory mechanisms have been realised in game-based contexts, and transparency of game design decisions and processes.

Complimenting the findings from Chapter 6, the fourth empirical chapter (Chapter 7) adopts an engaged research philosophy to contribute to the third research question. Working with AVIVA (a large multinational insurance organisation with offices in Norwich), this chapter draws from the findings of the previous chapters in the design and development of a new game-based intervention, *The disPHISHinformation Game*. After findings from the review highlighted that many game-based interventions did not comprehensively draw on inoculation literature, this study aims to take a more considered approach to incorporating inoculation theory, considering important inoculation mechanisms such as threat in its design. Moreover, this study also draws on instructive (serious) game design literature to guide the development of the intervention through interaction with The Transformational Framework (Culyba, 2018) and 'playcentric' development processes (Fullerton, 2004).

Finally, Chapter 8 concludes this thesis with a synthesis of empirical and theoretical contributions, summarising in turn how each research question has been addressed. The implications of this thesis's findings are discussed with reference to inoculation and misinformation literature, and an agenda for potential future research is outlined.

2. Literature Review

The risk that misinformation poses to democratic processes, public safety, and other population rights is continuously being demonstrated (Erizanu, 2024; M. Fisher et al., 2016; Roozenbeek et al., 2020), and the dangers of misinformation are becoming more prevalent (Ecker et al., 2024; Lewandowsky et al., 2017). Candidate solutions to tackling this pressing issue come from a wide selection of (often collaborating) fields, such as law, software engineering, politics, or psychology (Roozenbeek et al., 2023). Although misinformation only accounts for a small portion of news consumed by individuals (Allen et al., 2020; Guess et al., 2020b), total misinformation exposure remains high (European Commission, 2018; Neely et al., 2022). This issue demands attention from scientists and policy makers globally.

As misinformation has become more prevalent and garnered more attention, it has seen responses including new regulations (Regulation 2022/2065), media literacy campaigns (Wineburg et al., 2019), and psychological interventions (Roozenbeek & van der Linden, 2019). These often-overlapping fields have sought to address the risk misinformation presents; however, the frequency and manipulativeness of misinformation has changed as digital platforms and online social networks have evolved in an ever-changing digital landscape (Newman et al., 2024; Ofcom, 2024a). This chapter argues that individual-level interventions, such as inoculation theory, remain an accessible, engaging, and effective way to reducing susceptibility to misinformation, particularly when in a game-based format. This chapter aims to review core literature that underpins these interventions and introduce themes which form the basis of discussion and analysis later in this thesis.

In this chapter I first introduce the different schools of thought on solutions to fighting misinformation, including system- and individual-level approaches, and pre- and de-bunking. Secondly, I will present a central theory to this thesis, inoculation, discussing classical inoculation components and mechanisms, inoculation characteristics, and contemporary ideas around inoculation theory's use. Thirdly, I discuss serious games and how they can be used as an effective active learning method. I additionally present literature surrounding misinformation serious games using inoculation as a foundational theory, and discuss current open questions surrounding their design, and realisation of inoculation theory. Finally, I discuss some

contemporary conceptualisations of inoculation theory, making refence to existing literature on memory and learning.

2.1 Candidate Solutions to Fighting Misinformation

Solutions to fighting misinformation are diverse, often multidisciplinary, and attack misinformation at varying parts of the "toxic cycle of fake news" (Kozyreva et al., 2024; Parsons & Drünkler, 2019, p. 34). Candidate solutions range from legal regulations (e.g., the European Union's Digital Service Act, Regulation 2022/2065) to reviews of educational curriculum (UK Department for Education, 2024) and educational games (Hopkins et al., 2023). Work by Kozyreva et al. (2020) and Roozenbeek et al. (2023) categorises these diverse approaches to combatting the dissemination and effects of misinformation into two broad categories: system-level and individual-level. The classification of these approaches brings together frequently disparate research tackling (often multifaceted) misinformation. System-level approaches aim to tackle misinformation through its supply. Put simply, system-level approaches attempt to reduce the amount of misinformation disseminated, curbing the amount of misinformation reaching information consumers. Individual-level solutions on the other hand tackle the consumer-side of misinformation by changing individual behaviours or reducing susceptibility of individuals to misinformation. The different types of system-level and individual-level interventions are summarised in Table 1. The remainder of this section explores some of these approaches in turn.

Table 1: Different system-level and individual-level types of intervention against misinformation (Roozenbeek et al., 2023).

System-level Interventions			
Algorithms	**Business Models**	**Legislation**	**(Geo)politics**
• Transparency and accountability measures	• Addressing ad-tech • Supporting reliable news media	• Combatting online harms • Regulating tech platforms	• Combatting bad actors • Reducing polarisation
Individual-level Interventions			
Boosting	**Nudging**	**Debunking**	**Content labelling**
• Pre-bunking and inoculation • Critical thinking • Media literacy	• Accuracy primes • Social-norms nudges	• Fack-checking and corrections	• Automated or manual labels

System-level approaches to combating the spread of misinformation can be broadly categorised into changes to algorithms, legislation, (geo)politics, and business models. These system-level approaches share characteristics related to Lawrence Lessig's pathetic dot theory (Lessig, 1999), which stipulates that one's behaviour (e.g., susceptibility to misinformation) can be influenced by four main 'regulators': changes in law, influencing social norms (e.g., through education or advertising campaigns),

changes to market influences, and changes to architecture (technical infrastructure). These system-level approaches to fighting misinformation play an important role in improving online information environments, and some governments around the world have taken system-level actions to combat misinformation (Funke & Flamini, 2018).

Algorithms underpinning social media platforms such as finely tuned content processing and propagation algorithms determine the content users can post, see, and are recommended. These algorithms, especially content recommender algorithms, heavily influence the virtual information environment in which users exist. Confirmation bias suggests that users are more likely to select information that aligns with their personal beliefs or attitudes (Raymond S. Nickerson, 1998). Recommender algorithms are, to a certain degree, incentivised to prioritise and recommend content congruent with a user's beliefs and attitudes to maintain user satisfaction. This can lead to users existing in an 'echo chamber' of exclusively congruent (but not necessarily factual) information (Alfano et al., 2021; Del Vicario et al., 2016; Quattrociocchi et al., 2016), and can stimulate belief polarisation (Lucía et al., 2018; Vicario et al., 2016, cf. Bail et al., 2018). Furthermore, these recommender algorithms are rarely transparent, making research understanding the effects of algorithm design difficult (Roozenbeek et al., 2023). Adjusting recommender algorithms to consider potential side effects and providing access and resources to independent researchers could help to create a safer online information environment (Alfano et al., 2021; Roozenbeek et al., 2023). However, like many other system-level interventions, this requires willingness from content platforms to engage, or regulation to enforce algorithm changes.

Changing business models to, for example, move business incentives away from prioritising advertising revenue is another system-level approach to combatting misinformation (Rathje & van der Linden, 2023; Roozenbeek & van der Linden, 2024). One of the primary motivations of for-profit businesses is financial. In 2024 Meta recorded a revenue of 164,500 million dollars, 97.6% of which came from advertising (Meta Platforms Inc, 2024). Current business models provide strong business incentives to serve not only attention-grabbing content, but content that is more likely to be shared or reposted to a wider audience to maximise this advertising revenue (CCDH, 2024). Decisions by Meta to remove fact checkers on their platforms in favour of more open platforms (McMahon et al., 2025) and restarting a program that "paid bonuses to creators for content based on views and engagement" (Silverman, 2025, para. 5; "Earn money from", 2024) has supported this view. This is problematic, as the types of content that often drive high engagement include divisive content (Harris et al., 2023; Mousavi et al., 2022), negative content (Watson et al., 2024), content that is emotionally arousing (McCosker, 2013; McLoughlin et al., 2024; Rimé et al., 1991), and content that stimulates

outgroup animosity (Brady, Crockett, et al., 2020; Brady, Gantman, et al., 2020). Different business models such as subscription-based models, as Meta have offered in some European countries ("Facebook and Instagram to offer", 2024), could shift business incentives to prioritising user experience (Rathje et al., 2023), but this too relies on the willingness of social media platforms who may be naturally reluctant to change.

Legislation underpins many of the solutions proposed, both directly, and through enforcing other system-level solutions discussed. Many countries have created new laws to combat online harms (including misinformation), including the European Union Digital Services Act (Regulation 2022/2065) and to a lesser extent, the United Kingdom's Online Safety Act (Online Safety Act, 2023). These acts do not criminalise misinformation explicitly, but do so indirectly (i.e., through the regulation of digital services). The regulation of misinformation is difficult, however, as curbing the spread of misinformation has been used as a justification for reducing press freedom ("Censorious governments", 2021) and can raise questions of free speech (Gielow Jacobs, 2022; Nuñez, 2020). For example, during the COVID-19 pandemic Hungary passed legislation criminalising the spreading of misinformation. Many have argued that this law has been used to attack journalists and threaten press freedom (International Press Institute, 2020). Enforcement on malicious actors creating misinformation content is often difficult if they are in other countries that share different laws and values. Regulation of online platforms, such as the Digital Services Act, places responsibility on internet platform companies to maintain safe information environments. This too is difficult, however, as many of these platforms reside outside of the United Kingdom and, as such, can choose instead to simply restrict access to the platform in any country imposing regulation (as some have previously threatened, see Shead, 2022). Nonetheless, regulation against misinformation plays an important role in improving the online information environment (Roozenbeek & van der Linden, 2024).

Both national and international politics frequently plays a role in the spread of misinformation online. Social media platforms have become a major source for people to get news (Grieco, 2020; Ofcom, 2024a; Walker & Matsa, 2021), and political parties are now using these platforms in election campaigns as communication strategies (L. Fisher & Andringa, 2024; Fletcher, 2024; Moore, 2016). Indeed, political polarisation on social media platforms is inextricably linked to misinformation (Conover et al., 2021; Kubin & von Sikorski, 2021). Additionally, social media has been used by hostile foreign actors as platforms to spread disinformation with the aim of achieving (often political) malicious objectives (Erlich & Garner, 2023; McGeehan, 2018; Morley-Davies, 2024; Roozenbeek, 2024). When assessing system-level solutions to fighting misinformation, addressing these (geo)political issues may help to attack misinformation from a prominent source. These (and other) approaches to reducing

misinformation play an important role in improving the online information environment, but many of them centre around reactive (rather than proactive) legislation, and online platform alterations whose implementation remains challenging due to differing business incentives (Blair et al., 2024; Parsons & Drünkler, 2019).

Individual-level approaches focus on the information consumer through building skills, awareness, or cognitive resilience to misinformation. Individual-level approaches have received particular attention in research as they give individual people autonomy to improve their own news hygiene. Roozenbeek et al. (2023) categorise individual-level interventions into content labelling, nudging, debunking, and boosting.

Content labelling includes any practice in which content is (either automatically or manually) labelled with information on its veracity or reliability. This approach, which includes credibility and trustworthiness indicators and crowdsourced fact-checking, has been found to be effective (Allen et al., 2021; Celadin et al., 2023; Pennycook & Rand, 2019; Sethi, 2017; Thorne & Vlachos, 2018; Yaqub et al., 2020). Automated content labelling has been an approach adopted by some social media platforms as this requires little human intervention and can be applied quickly at scale (Roozenbeek et al., 2023). Labelling algorithms are not infallible, however, and a small percentage of incorrect content labelling can be costly when scaled up (Alaphilippe et al., 2019). Moreover, the effectiveness of content labelling is not assured, and some research has questioned the effectiveness of this approach (Brandtzaeg et al., 2018; Kelly Garrett & Weeks, 2013; Oeldorf-Hirsch et al., 2020), with some experiments observing boomerang effects (Nyhan & Reifler, 2010). Fact checking is a subset of content labelling, and until recently an adopted practice by many social media platforms (McMahon et al., 2025; Silverman, 2025). It too has been found to be effective (Martel & Rand, 2024), however, like many system-level interventions, requires willingness from social media platforms to adopt.

Nudging, another individual-level intervention that overlaps with content labelling, involves changes to platforms that affect user behaviour without restricting options or functionality available (Fazio, 2020; Pennycook et al., 2020). Where content labelling may include the removal of content (e.g., through fact checking), nudging does not restrict the user's options. Nudging outcomes can be difficult to replicate in real-world environments (versus lab-based research), however (DellaVigna & Linos, 2022), and some evidence suggests that their effectiveness may not be long lasting (Sasaki et al., 2021). Moreover, like content labelling and many system-level interventions, these solutions often rely on either regulation, or willingness of platforms to implement, which is not guaranteed (McMahon et al., 2025). Boosting the skills and cognitive resilience to misinformation may reduce the spread of misinformation without

reliance on the willingness of social media platforms to implement experimental content labelling or nudging systems.

Two popular individual-level approaches to tackling misinformation are pre-bunking and debunking. Debunking aims to correct misinformation claims after they have been consumed by an individual. A challenge associated with debunking is individuals are not always cognitively motivated to change their (misinformed) belief, particularly if it is congruent with other held beliefs (e.g., politically polarising topics. See Chan & Albarracín, 2023; Ecker et al., 2011). Pre-bunking, conversely, is a boosting approach to tackling misinformation that helps individuals become better in their own decision-making competence (Kozyreva et al., 2024). By warning an individual of false arguments they may face in the future or pre-emptively training them on persuasion techniques used in misinformation, individuals can build resistance to future attacks on their beliefs. One approach to pre-bunking that has received particular academic attention is inoculation theory.

2.2 Inoculation Theory

Originally explored by William J. McGuire in the 1960s (McGuire, 1964; McGuire & Papageorgis, 1962; McGuire, William J. Papageorgis, 1961), inoculation theory is an approach to building resistance to persuasion. McGuire reflected that people characteristically defend held beliefs from counter-attitudinal arguments through avoidance. He speculated that – just as an individual who has been raised in an antiseptic environment may be more vulnerable to disease (Bloomfield et al., 2006) – defending one's beliefs by avoiding contrasting information may leave a person poorly prepared to defend their convictions if unwillingly exposed to counter-attitudinal arguments (McGuire, William J. Papageorgis, 1961). The theory follows the metaphor of a biological vaccine, positing that by being pre-emptively exposed to a small, altered 'dose' of an attack to a belief, resistance can be built to defend that belief from future persuasive attack. Despite its early inception, inoculation theory has enjoyed continued interest from the research community, with the theory evolving and becoming better understood over time. Meta-analysis research has shown that inoculation theory is generally an effective approach to building resistance to persuasion (Banas & Rains, 2010).

In McGuire's original vision of inoculation theory, the key mechanism to a successful intervention was threat: the individual must be motivated to develop cognitive defences against an incoming attack, which was thought to be better motivated by threatening the prior belief rather than reassuring it (McGuire, 1964). McGuire posited that an effective way of delivering threat is by forewarning an individual of their attitudinal vulnerability, and pre-emptively refuting arguments

against it. The individual is therefore better able to defensively counterargue against subsequent attacks to their belief and is motivated to continue to do so. These two key mechanisms, threat and counterarguing, have continued to dominate the inoculation literature and have been the focus of much research as subsequent scholars have attempted to disentangle them and understand the role they play in conferring resistance (Compton, 2024; Compton & Pfau, 2005; Pfau et al., 1997).

In conventional inoculation theory research and application, these mechanisms have been realised through two distinct components. First, an explicit forewarning can deliver threat to the reader, informing the individual of their attitudinal vulnerability (see for example M. L. M. Wood, 2007). Secondly, refutational pre-emptive messages expose individuals to weakened forms of counter-attitudinal arguments and can provide direct refutations to these arguments, which the individual can use to defend themselves in subsequent attacks. As well as refuting specific arguments individuals may come across in subsequent attitudinal attacks, pre-emptive messages are seen to improve counterarguing skills and demonstrate how to effectively engage in attitudinal defence through practice (Ivanov et al., 2016). Furthermore, these messages can provide a covert form of attitudinal threat (by rendering the real threat, see McGuire, 1964). This process is visualised in Figure 1.

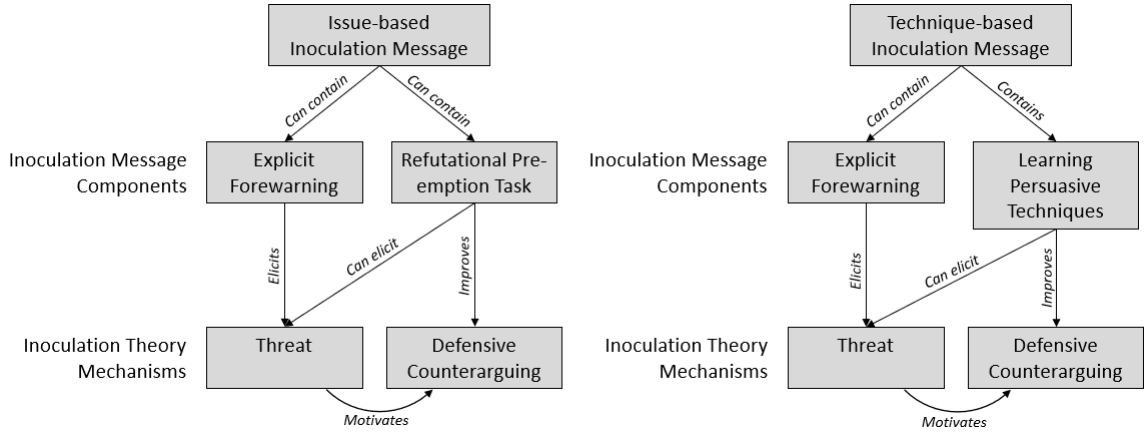


Figure 1: Breaking down McGuire's theory into its core cognitive mechanisms, and typical message components often used to stimulate these mechanisms.

McGuire saw threat as conceptually important to inoculation theory, positing that rather than bolster a prior belief, an intervention should threaten it, motivating the receiver to build cognitive defences to future attack (McGuire, 1964). McGuire himself never actually measured threat, but other scholars have subsequently sought to measure its mediating effect on conferred resistance and have found conflicting results (e.g., Banas & Rains, 2010; Banas & Richards, 2017; Compton & Pfau, 2005). Traditionally, threat has been measured as being related to fear (i.e., measuring if the individual feels literally threatened), and a scale developed by Burgoon et al. (1978) that measures this apprehensive form of threat has been commonly used in inoculation

research over the years (Banas & Richards, 2017; Compton, 2021). Studies have suggested that apprehensive threat does not have a direct relationship with resistance, but instead increases issue involvement, which can act as a mediator between threat and resistance (Compton & Pfau, 2004). Unintuitively, however, it does not appear that higher levels of apprehensive threat confer more resistance. Instead, apprehensive threat appears to act more like a threshold, with a certain level required for a successful inoculation (Banas & Rains, 2010; Compton, 2021). Importantly, however, McGuire's original conceptualisation of threat was closer to a recognition of attitudinal vulnerability (rather than a literal feeling of threat), and some scholars have more recently called for threat to be reconsidered as being more motivational in nature (Banas & Richards, 2017; Compton & Ivanov, 2012).

As scholars have sought to better understand and apply the internal mechanisms of inoculation theory over its relatively long history, its potential to offer protection broader than any individual issue-space has been revealed. Early work by McGuire was limited to inoculating individuals against attacks on widely accepted beliefs on single issues, such as teeth-brushing and the benefits of penicillin (McGuire & Papageorgis, 1961). In these contexts, refutational pre-emptive message components introduced and refuted arguments specific to these topics. Since then, issue-based inoculation has been applied to a wide range of more controversial issues (e.g., Pfau et al., 2008). Importantly, research has demonstrated that conferred resistance can carry over to different arguments on the same topic (Banas & Rains, 2010), or even different (but related) topics entirely (Parker et al., 2012, 2016). To further the scalability of inoculation theory, scholars have more recently sought to inoculate against persuasion techniques observed in misinformation, rather than (fallacious) arguments on a single issue (Cook et al., 2017; Lim & Ki, 2007; Roozenbeek & van der Linden, 2018, 2019). This so-called technique-based inoculation focuses on the rhetorical techniques found in misinformation (see Figure 1) and represents an actualisation of inoculation theory for many recent misinformation interventions (Compton, 2024). By moving away from issue-based inoculation and instead focussing on techniques, interventions have been found to confer resistance effectively on a wide range of content that makes use of these techniques (Lewandowsky & van der Linden, 2021), representing a 'broad-spectrum vaccine' (Roozenbeek & van der Linden, 2019). Some authors have made a distinction between technique-based inoculation and logic-based inoculation, highlighting that logic-based approaches focus on logical fallacies over rhetorical techniques (Hopkins et al., 2023). However, the distinction between logic-based and technique-based interventions is contested (Roozenbeek & van der Linden, 2024), suggesting a lack of clarity in this area of the literature.

As applied inoculation research has increasingly taken place outside the lab in real-world contexts, researchers have had to manage the noisy environments in which participants experience and learn new information. Frustratingly, participants taking part in this research are in many cases being inoculated after they have already been subjected to the attack. McGuire's initial work focussed on a lab-condition environment, where participants were inoculated before an attack, analogously parallel with a prophylactic vaccine. This is rarely the case in the real world; one study found, for example, that of a sample of full-time working professionals in the US, over 50% reported to have been exposed to conspiratorial vaccine misinformation (Lee et al., 2022). Scholars have therefore begun investigating whether inoculation interventions delivered after an attack still confer protection (Ivanov et al., 2016). Research has found therapeutic interventions to still be effective (Ivanov et al., 2016). Research by Ivanov et al. (2022) has suggested that threat and counterarguing play a lesser role in therapeutic interventions, but how this may inform future intervention design remains to be seen. Important to note, is a lack of clarity around what it means for an intervention to be 'therapeutic'. Communication scholars have defined 'therapeutic' as describing an intervention where the individual is already 'infected' (i.e., starts with an opposing attitude, see Compton, 2020). However, previous game-based inoculation research describes therapeutic inoculation as where the individual has been previously exposed (but not necessarily 'infected', see Basol et al., 2021; Roozenbeek & van der Linden, 2019). This thesis adopts the broader definition (including both exposure and infection), but I highlight these contradicting definitions here for clarity.

Some scholars have begun to explore the potential for conferred resistance to misinformation to spread through word-of-mouth communication. Inoculation has conventionally been seen as a subvocal process. However, post-inoculation talk (PIT) may be a platform for positive effects to transfer between individuals as they look to advocate their position with others (Compton, 2024; Compton & Pfau, 2009). An advantage of the inoculation theory analogy is that investigations in boundary conditions are often more accessible as a result of their being analogous with medicine. This too is the case with PIT, as scholars try to understand if 'immunity' can spread from inoculated individuals to those not inoculated through in-the-field social interaction (Compton & Pfau, 2004; Ivanov et al., 2012).

Work equally inspired by the inoculation analogy has found benefit in follow-up exercises, or 'booster sessions' (Capewell et al., 2023; Maertens et al., 2021, 2025). Research by Maertens et al. (2025) has shown that by readministering an intervention to participants, conferred resistance can be bolstered. This is also the case for shorter interventions, or even simply reminding individuals of the intervention through empirical measures (e.g., Maertens et al., 2021). Considering 'replayability' for

inoculation interventions may lead to interventions that can be used as 'boosters', however intervention attractiveness, uptake, or retention is rarely tested explicitly, and this is therefore rarely considered in design (Davies et al., 2024; A. Johnson & Madsen, 2024; Roozenbeek et al., 2024).

Feedback tasks are another contemporary inoculation theory component that may create more effective interventions. Recent research in inoculation theory has found that debriefing, often a part of the study methodology rather than the inoculation message itself, can have a positive immediate effect on individuals, and can help reduce decay of conferred resistance. Leder et al. (2024) argues the importance of immediate feedback exercises in recent research, although the significance of 'debriefing' sessions for experiential learning has been identified before (Pearson & Smith, 1985). Most research integrates some form of debrief, although this is often primarily a requirement from ethics boards. Although we don't yet have empirical evidence for the benefit of debrief sessions in gamified contexts, work by Leder et al. (2024) and Maqsood and Chiasson (2021) is promising and suggests that there may be benefits to considering debrief/feedback in game design.

Although initial research by McGuire focussed on simple reading and writing tasks, scholars and practitioners have since delivered inoculation messages through increasingly active and innovative means. In McGuire's original research, two different types of pre-emptive refutation were tested: a passive intervention that involved reading counterarguments, and an active intervention that involved participants writing their own counterarguments. Active inoculation is commonly seen as being more effective due to the cognitive process being more involved (M. Green et al., 2022), but research has found evidence favouring both approaches (Compton & Pfau, 2005). Researchers have more recently begun to leverage the potential benefits of serious games as a platform to host active inoculation interventions (Cook et al., 2023; Jeon et al., 2021; Roozenbeek & van der Linden, 2019). It is argued that serious games can offer not only a more active platform for inoculation interventions than conventional writing tasks (Roozenbeek & van der Linden, 2018), but by inoculating on misinformation techniques, games can reach wider audiences offering better scalability than conventional inoculation messages (Roozenbeek & van der Linden, 2019). The first applications of inoculation theory to gamified contexts were exploratory studies by Roozenbeek and van der Linden (2018) with the collaborative board game *Fake News*, an analogue card game to be played with groups of four. This led to a large-scale study with the digital game *Bad News* (Roozenbeek & van der Linden, 2019), from which the field has slowly grown.

2.3 Serious Games and Misinformation

Serious games (also known as applied games or transformational games) sit at the intersection between games and education and are becoming an increasingly popular form of active learning in contemporary classrooms and misinformation intervention research (Kiili et al., 2024; Zhonggen, 2019). This section introduces important literature on serious games relevant to this thesis, including conceptual and empirical benefits of serious games, knowledge gaps around serious games as educational approaches, and the application of inoculation theory in games-based contexts.

Serious games are defined as complete, purposeful games that are designed to meet objectives beyond entertainment (Abt, 1970). Although the terms are often used interchangeably, serious games differ from gamification, which is the application of game elements (such as points scoring and competition with other players) to non-game contexts (Deterding et al., 2015). Instead of entertainment, the core objective of serious games is education. This can be in the sense of conventional educational objectives (for example, the game *The Oregon Trail* educated players on pioneer life on the Oregon Trail), improving skills (simulators such as *X-Plane* can help build skills in a game-based context, see also Hill et al., 2006), or strategic planning (serious games have long played a part in military contexts as 'wargames', see Curry & Drage, 2020).

Serious games offer attractive benefits for practitioners as educational tools. Notably, serious games are commonly asserted to a) increase learning gains in players over conventional learning methods, b) increase learner motivation to interact with educational materials (through the serious game, or otherwise), and c) yield higher knowledge retention than conventional methods (Gee, 2003; Wouters et al., 2013). Some empirical studies with serious games have supported these claims (Barab et al., 2009; Bengston et al., 2022; Cameron & Dwyer, 2005; Jarvis & De Freitas, 2009), and have additionally found that serious games can also elicit positive affect both in the player (Lau et al., 2017), and towards learning (Lamb et al., 2018). Serious game effectiveness has also been observed to increase when combined with other instructional methods, such as discussion and explicit practice (Connolly et al., 2012a; Wouters & Van Oostendorp, 2013; Yaman et al., 2008). The active nature of serious games is also supported by ideas in educational psychology, which suggest that "active cognitive processing of educational material is a prerequisite for effective and sustainable learning" (Wouters et al., 2013, p. 250).

In contrast, some scholars have struggled to measure these increases in motivation to engage with learning materials (Lorant-Royer et al., 2008; Stanfield & Connolly, 2008; Wouters & Van Oostendorp, 2013), and one meta-analysis of serious games as educational tools suggests that there are in fact no significant benefits over

traditional materials (Girard et al., 2013). A number of systematic reviews have been completed of serious games in different contexts, including education, healthcare, and smaller topics such as cyber bullying and mental health. Reviews suggest that although the number of serious games in research is abundant (Zhonggen, 2019), there is a lack of empirical research investigating the effectiveness of serious games in learning over traditional methods (Damaševičius et al., 2023; Girard et al., 2013; Lau et al., 2017; Mesa-Gresa et al., 2018). More research is required to compare different types of learning and understand how players retain knowledge from these types differently (Girard et al., 2013; Lorant-Royer et al., 2010). Furthermore, research examining how specific serious game design decisions result in better learning could lead to more effective serious games (Boyle et al., 2016).

Despite these uncertainties surrounding their effectiveness, scholars and practitioners have applied serious games in education and research as a tool to actively inform players on the techniques and risks of misinformation. One early serious game educating players on journalism and misinformation is *BBC iReporter* (Cellan-Jones, 2018). Players take the role of journalists and are faced with a fictional major breaking story related to a cyber security incident. Players must make journalistic decisions on what to publish, making trade-offs between, for example, evaluation speed and information veracity. Other misinformation serious games take different player perspectives: *Factitious* (Grace & Hone, 2019) and *MAthE* (Katsaounidou et al., 2019) ask players to label news items as being true or false, giving players feedback after their choices, with some positive preliminary results (Katsaounidou et al., 2019).

Researchers have more recently begun to leverage the potential benefits of serious games as a platform to host active inoculation interventions. McGuire's initial research was primarily lab-based, and efficacy of the theory was tested with variables including type of defence, and the level of participation in the intervention (McGuire & Papageorgis, 1962; McGuire, William J. Papageorgis, 1961). Much subsequent research focussed primarily on passive interventions (Roozenbeek & van der Linden, 2018); however, games have more recently offered an alternative approach to employing active involvement. It is argued that serious games offer not only a more active platform for inoculation interventions than conventional writing tasks (Roozenbeek & van der Linden, 2018), but by inoculating on misinformation techniques, games can reach wider audiences offering better scalability than conventional inoculation messages (Roozenbeek & van der Linden, 2019).

Serious games using inoculation as a foundational theory have been applied to a variety of topics such as scams (Roy et al., 2025), social engineering (Aladawy et al., 2018), and misinformation (Compton et al., 2021). The card game *Fake News* (Roozenbeek & van der Linden, 2018) was one of the first serious games to inoculate

against misinformation, and asked players to work collaboratively to create 'fake news' articles in the style of a given character. Scholars have subsequently applied game-based interventions to sub-topics of misinformation, including misinformation in politics (Roozenbeek & van der Linden, 2020), health (Basol et al., 2021), climate change (Cook, 2021), multimodal misinformation (Neylan et al., 2023), echo chambers (Jeon et al., 2021), and more (DeJong, 2023; Devasia & Lee, 2024; N. Henderson & Pallett, 2024; Kiili et al., 2024). Despite the growing popularity of game-based misinformation interventions, there remain questions on their design and efficacy that require further attention (Roozenbeek et al., 2024).

Despite the growing number of misinformation games using inoculation as a foundational theory, most of these have been digital interventions citing scalability and ease of testing as justification. Games such as *Bad News* have been the focus of large-scale global studies, allowing an interrogation of the role of memory (Maertens et al., 2021), different forms of misinformation (Neylan et al., 2023; Roozenbeek et al., 2022), and further elements such as message components (Maertens et al., 2025). This has resulted in comparatively little research in analogue and virtual reality serious games (DeJong, 2023; Kiili et al., 2024). The potential for the characteristics of alternative platforms to differently moderate conferred resistance should not be ignored, and some serious game designers have suggested that social features of analogue games may help to better stimulate learning outcomes in some environments (Graffer et al., 2015; N. Henderson et al., 2024). Demographic breakdowns of player bases for analogue games also reveal their popularity for not only a wider range of ages, but also a large neurodiverse community (Cross et al., 2023).

As well as a homogeneity of game platforms, very few interventions have assessed the long-term effect of interventions to understand how inoculation effects decay. Understanding how long the effects of an inoculation intervention last are important, as a high decay may signal a need for 'booster' interventions (Maertens et al., 2025), or alternative approaches entirely. The real-world potential for inoculation interventions to build resistance to misinformation has been hampered by our limited knowledge in intervention longevity (N. Henderson et al., 2023; Maertens et al., 2021, 2025). Current research suggests that the decay of intervention effects may follow a forgetting curve, and 'booster' interventions could bolster conferred resistance in participants (Maertens et al., 2021, 2025). However, most longitudinal research with game-based interventions have been limited to the digital game *Bad News*. Moreover, it is the author's understanding that there is currently no longitudinal research with any analogue interventions. This is of note as tactile game elements and the tangible and embodied experience of an analogue serious game may reduce intervention decay

differently than digital interventions (Cutumisu et al., 2019; Edwards et al., 2024; N. Henderson & Pallett, 2024).

Despite many misinformation games using inoculation as a foundational theory have shown to be efficacious, success has not been absolute, and scholars have identified some open questions that require further investigation (Compton, 2024). As well as knowledge gaps in analogue design and longitudinal decay (Maertens et al., 2021), the field may naturally suffer from publication bias (Banas & Rains, 2010; Edwards et al., 2024; Lamb et al., 2018). Researchers have also expressed uncertainty around 'what works' in intervention design (Appel et al., 2025). Research by Harjani et al. (2023) found that their digital game-based intervention *Join This Group* had no significant effect on players in northern India and have suggested that further research is needed in adapting game-based interventions for target audiences. Despite these limitations and knowledge gaps, game-based interventions remain a promising approach to building resistance to misinformation in players.

2.4 Considering Inoculation Through the Lens of Learning

As inoculation theory has received continued interest over the years, key mechanisms have seen scrutiny as communication scholars have sought to clarify and better our understanding of them. This section discusses the more recent novel conceptualisation of the inoculation processes: one of learning and forgetting. To compliment, this section additionally introduces literature on experiential and collaborative learning.

The role of threat, for example, has repeatedly been identified as needing further investigation (Banas & Rains, 2010; Banas & Richards, 2017; Compton, 2009, 2021; Ivanov, 2017; Pfau et al., 2010). Scholars have often struggled to identify apprehensive threat as a moderating variable (i.e., a variable that affects the extent to which an intervention confers resistance, Banas & Rains, 2010), and some have struggled to find evidence for traditional threat's moderating effect to resistance at all (Compton, 2009). Understanding the role of threat in inoculation theory interventions is important as, if threat plays a lesser role than other inoculation mechanisms, this has major implications for intervention design. Some scholars have suggested that delivered threat (whether through a distinct forewarning or implicitly in refutational pre-emption) should be considered more motivational in nature (Banas & Richards, 2017), acting as a 'catalyst' to participate in the inoculation process, rather than a response of apprehension (Banas & Richards, 2017; Compton, 2021; Maertens et al., 2025).

The role of threat as a source of motivation to build resilience to misinformation has also been questioned in longitudinal research. Work by Maertens

et al. (2021) and Maertens et al. (2025) has sought to understand how effects from game-based interventions decay, finding that the decline of resistance over time may be better explained by memory and forgetting rather than a decrease in threat. When measuring threat, motivation, and memory in participants, Maertens et al. find that "memory is one of the most dominant factors in intervention success and longevity" (Maertens et al., 2025, p. 9). This relationship between inoculation theory and memory and learning is explained in Maertens et al.'s 'memory-motivation' model (see Figure 2). The memory-motivation model does not negate the role of threat; instead, threat is conceptualised to motivate the participant to participate in the memory creation process, thus resulting in a more effective intervention. This recent new conception of inoculation theory – one of learning and forgetting – is novel and marks a potentially impactful shift in inoculation literature, introducing insights from learning pedagogy and memory research. Moreover, these findings have implications for practitioners who may seek to introduce concepts from pedagogy and memory research to make more memorable interventions that confer longer lasting resistance (Capewell et al., 2024; Leder et al., 2024; Maertens et al., 2025). However, this proposed model of inoculation-conferred resistance raises definitional questions about what is classed as an inoculation intervention, versus other communication strategies such as two-sided messaging.

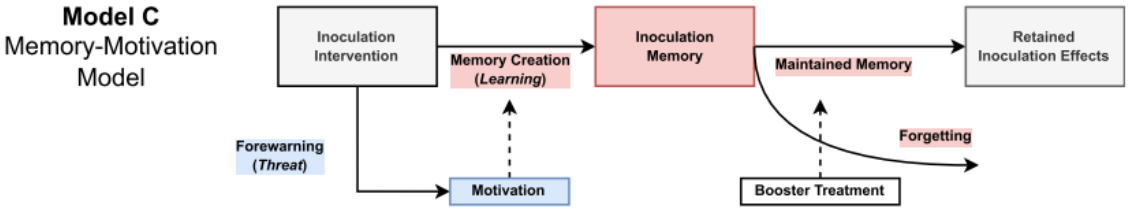


Figure 2: The memory-motivation model of inoculation theory (Maertens et al., 2025).

Pedagogy research plays a role in helping practitioners create effective learning content but is largely unexplored in inoculation literature. Although scholars have previously discussed the role learning may play in inoculation (Pfau et al., 2005; Maertens et al., 2001), the wide and rich literature on learning and learning theories has yet to be explored. Learning theorists have been proposing different models for how people learn since the early 20th Century (Dewey, 1938), but these models (e.g., experiential, reflective, or collaborative learning) have not yet been drawn upon in an inoculation context. Learning theories allow the modelling of how people acquire new knowledge and can be a powerful tool for practitioners. Instead of behaviourism, where learners are considered in terms of being shaped by external stimuli and responses, we consider the interventions used and discussed in this thesis to be more constructivist in nature: recognising the player's active role in learning through experience and interaction with their environment (Piaget, 1972). Specifically, this thesis engages with

work by Dewey (1938), Piaget (1972), Kolb (1984), D. W. Johnson & Johnson (1991), Schön (1984), and Bandura (1977) through educational theories such as experiential learning, collaborative learning, reflective learning, and social learning.

Constructivist educational theorist Piaget (1972) believed that people learn through experience. This prominent role of experience in learning has prompted experiential learning models from Kurt Lewin, Dewey (1938), and particularly Kolb (1984), who believe that knowledge is created through the transformation of experience. Although differences exist between models, they share four core stages: an initial experience, observations and reflections, formation of abstract concepts, and testing implications of concepts in new situations (Kolb, 1984). Put simply, experiential learning posits that from a 'concrete experience' (such as a game-based misinformation intervention), one can reflect, develop new concepts and theories from said reflections, and apply this new knowledge in new situations (Kolb, 1984; Kolb & Fry, 1975). Serious games are applications of experiential learning, as they offer not only a concrete experience, but provide feedback to players in a similarly iterative learning process (Boyle et al., 2011; Connolly et al., 2012b).

Educational theorists including Dewey (1933), Schön (1984), Boud et al. (1985), and Boyd and Fales (1983) explored the reflective elements of learning from experience, focussing on different forms of reflection (e.g., reflection-in-action and reflection-on-action, see Schön, 1984). Much of the research investigating reflective learning focusses on the types of reflection occurring on experience, and how practitioners such as educators and trainers can increase the meaning and value of reflective practices (Scanlan & Chernomas, 1997; Schön, 1984; Zeichner, 1994). Reflective learning can be leveraged through a number of different mediums (K. Henderson et al., 2004) including journaling (Bain et al., 1999; National Health Service, n.d.), peer and self-assessments (Boud, 1990), fieldwork diaries (Fuller et al., 2006), reflective commentaries (Maclellan, 1999), and action research (Avison et al., 1999).

Collaborative learning is an application of constructivist-like principles to group contexts, positing that group-based activities can be powerful approaches to learning (Dillenbourg, 1999; D. W. Johnson & Johnson, 1991, 1987; Keyser, 2000). D. W. Johnson and Johnson (1987) suggest that practitioners can structure learning to be competitive, individualistic, or cooperative, and suggests that cooperative learning has been a largely under-utilised educational approach. Cooperation between players is also an increasingly popular design approach in games (Bayeck, 2020; Peppler et al., 2013; Rogerson & Gibbs, 2018), more recently in online settings with the improvement of internet speed and infrastructure (Zagal et al., 2006). This is the case for serious games as well, with scholars finding positive outcomes for collaborative serious game designs (Johnson-Glenberg et al., 2014; Laine & Lindberg, 2020; Peppler et al., 2013;

Wendel et al., 2013). Although most game-based inoculation interventions against misinformation are single player (N. Henderson & Pallett, 2024), some interventions allow players to work cooperatively to achieve a common goal (Adams, 2023; Pomichal & Trnka, 2022, 2023; Roozenbeek & van der Linden, 2018).

With its roots in behaviourism (rather than constructivism), social learning posits that one can acquire knowledge and skills through observations of others (Bandura, 1977). Put simply; by observing the consequences of actions by others, one can make better informed choices. Although initial research focussed on largely non-verbal observation, social learning theory can be applied to verbal communication too. Social cognitive theory dictates that people are not only influenced by observations of their environment, but they also influence the environment which they are in (Bandura, 1986). An instructive application of this is collaborative learning, where individuals learn from one another while collaboratively completing a task. Both theories apply to games, where players can be influenced by others' interactions with the game, and mechanics that induce challenge and feedback (Bandura & Cervone, 1983; Booker & Mitchell, 2021). Social learning has in the past been drawn upon in inoculation message design (Perry et al., 1980; Wallack & Corbett, 1987), although some scholars have criticised this hybrid approach for not considering the inoculation threat mechanism (Compton & Pfau, 2005).

Learning theories have been used to both describe the nature of learning occurring in serious games and inform the design of serious games themselves, but they have seen little application to inoculation theory. Serious games, and in particular simulation games (i.e., games that simulate real world activities), are by nature experiential (Gee, 2003) and offer players a concrete experience on which to reflect, often integrating feedback mechanics in an iterative learning process (Cowley et al., 2013; Marsh, 2016). Multiplayer serious game design can also facilitate cooperative learning between players (Echeverría et al., 2011; Peppler et al., 2013; Zagal et al., 2006), and scholars have suggested that competition can positively influence collaboration and collaborative learning, providing a social motivation for 'productivity' (Deutsch, 1949; D. W. Johnson & Johnson, 1991; P. Williams & Sheridan, 2010). Conversely, coaction between inoculation theory and learning theories has been uncommon. Some very limited work has explored the integration of Bandura's social learning theory in inoculation theory interventions (Perry et al., 1980; Wallack & Corbett, 1987), but other learning theories remain broadly unexplored.

Investigation into the boundary conditions of inoculation theory has at times raised questions on what truly classes as an 'inoculation theory' intervention. Threat, for example, is the central mechanism to inoculation theory (Compton, 2024; McGuire, 1964), and approaches that seek to integrate elements of inoculation with learning

theories have not always recognised this threat mechanism (such interventions could be considered closer to two-sided messaging, see Compton & Pfau, 2005; Lumsdaine & Janis, 1953; Wallack & Corbett, 1987). The threat mechanism itself has seen evolution over time, with scholars more recently proposing a 'motivational' interpretation of delivered threat (Banas & Rains, 2010; Banas & Richards, 2017). Other elements of inoculation theory have also raised questions of taxonomy; therapeutic interventions (see Section 2.2) have raised questions as (a) such intervention may be better described as persuasion rather than inoculation (Compton, 2012), and (b) mechanisms like threat may operate differently if participants' 'vulnerable position' is already compromised (Compton, 2020). The epidemiologic analogy on which inoculation theory is based may accommodate much of this contemporary inoculation research, however, and some scholars have argued that the inoculation analogy should be "more instructive than restrictive" (Compton, 2012, p. 233).

Conclusion and Research Questions

This chapter has presented the scientific literature from which this thesis is motivated and introduced foundational concepts from which empirical chapters are built on. Specifically, I have summarised system- and individual-level candidate solutions to reducing misinformation through the lens of frameworks by Kozyreva et al. (2020) and Roozenbeek et al. (2023). I have discussed inoculation theory as a strategy to build resistance to persuasion, discussed how inoculation has evolved over time, and how inoculation has been applied to misinformation. I have discussed how serious games combine the motivational affordances of games to deliver educational content, and their application to inoculation theory. Finally, I have introduced several key learning theories, including experiential, collaborative, reflective, and social learning, and discussed how these relate to serious games and inoculation theory.

Although changes to algorithms, business models, and legislation may be particularly effective approaches to reducing misinformation, these often require willingness from content platforms such as Meta and X to make foundational changes that may have negative financial impacts, or (in the case of legislation) are slow and often reactive rather than proactive. I suggest that individual-level pre-emptive approaches that help information consumers become more resilient to misinformation (such as inoculation theory) may be a more accessible approach to fighting false and manipulative information. Specifically, this literature review proposes that serious games (informed through learning theories) may be an effective platform to host inoculation interventions.

This literature review also draws attention to research gaps within the context of game-based inoculation interventions against misinformation. With relation to

inoculation theory's application to misinformation, this literature review highlights the lack of longitudinal research, with little known on (a) how conferred resistance lasts over time, and (b) how interventions may be better designed to last longer. Looking at game-based inoculation theory, it highlights the overreliance on quantitative methods, with very little work understanding why interventions are (in)effective (Roozenbeek et al., 2024). Furthermore, the chapter reviewed how most serious games against misinformation have been digital, with any potential collaborative or social advantages of analogue interventions going largely underexplored (DeJong, 2023; Kiili et al., 2024). Finally, I have highlighted a lack of coaction integrating learning theory research into inoculation intervention design, whose learning theory models may help practitioners design more effective inoculation interventions.

Informed by this literature review I now propose three research questions which drive the motivation and research design of this thesis. These have been defined through the research justification and motivation discussed above. Through attempting to answer these research questions, I hope to contribute knowledge to game-based inoculation literature, demonstrate novel approaches to game-based inoculation research and intervention design, and inspire future inoculation research.

RQ1. How effective are analogue game-based interventions against misinformation, and to what extent does this change over time?

RQ2. What kinds of learning are stimulated by analogue game-based inoculation interventions against misinformation?

RQ3. How could game-based inoculation interventions against misinformation be better designed?

3. Research Methods

This chapter provides an overview of the methodology adopted in this thesis. I outline and justify the research philosophy adopted, present and justify the research design for the three empirical studies of the thesis and highlight how findings from each study helped inform the design of the next. First, this chapter discusses how adopting an inductive pragmatist research paradigm enabled a flexible, mixed-methods approach to addressing the three research questions. Following this, this chapter addresses the three studies of this thesis individually, giving a comprehensive overview of the research methodology for a) the mixed-methods learning study with the game-based intervention, b) the review of game-based inoculation interventions against misinformation, and c) the design and development of the social engineering cyber security intervention *The disPHISHinformation Game*. This chapter also justifies the research design and methods adopted for each study in relation to the research questions while maintaining transparency surrounding the limitations of each study.

Research Approach

Building from theories of inoculation and broader misinformation research, many existing studies exploring the outcomes of game-based inoculation interventions have focussed on whether interventions improve players' misinformation susceptibility or other quantitative markers (e.g., Harjani et al., 2023; Roozenbeek & van der Linden, 2018, 2019). Much of this research is deductive in nature, confirming or rejecting hypotheses on intervention efficacy deduced from existing theory and literature. The research questions addressed in this thesis, however, demand that the present research goes beyond asking if an intervention is successful, but also *why*. To this end, this thesis takes a more inductive stance; in addressing these research questions I hope that the implications of this thesis are that "the findings are fed back into the stock of theory, and the research findings associated with a certain domain of enquiry" (Bryman, 2012, p. 24). Through answering these research questions, the theories and practices of game-based inoculation design can be improved for practitioners.

To this end, this thesis also adopts a pragmatist research paradigm. As highlighted in the comments above, existing research with game-based inoculation research has followed a positivist paradigm, relying on quantitative methods to determine intervention efficacy. Scholars have, however, expressed frustration at not understanding 'why' interventions are (in)effective (Appel et al., 2025; Harjani et al.,

2023; Roozenbeek et al., 2024; Saleh et al., 2024). Quantitative measures of intervention effectiveness do not tell the whole story: a deeper exploration of the personal and contextual experiences players have with interventions in required. Indeed, the literature review in the previous chapter has highlighted that the collaborative or social advantages of analogue game-based design have been underexplored. Therefore, the methods in this thesis have been chosen for their applicability to the research questions defined. The pragmatist research paradigm adopted in this thesis have been mixed methods in nature, including quantitative surveys, session observations, and semi-structured interviews.

Finally, this thesis has adopted a consecutive research strategy to designing individual studies. Although the design methodologies of the three studies presented in this thesis are introduced in advance of the chapters in which their findings are presented and discussed, these studies were not designed together. Instead, a data-driven design approach was taken; the findings of each study presented in this thesis directly informed the design of the next, allowing for a deeper interrogation of open questions raised. This represents a more engaged research approach. Through interacting with participants through qualitative methods, this research considers participants as partners rather than subjects.

Having introduced the research philosophy and paradigms adopted, I now move to introduce the individual studies presented in this thesis. Each study has been designed to address one or more research questions outlined above. The first study, a mixed-methods study with the analogue game-based intervention *Fake News*, seeks to answer the first research question. The second study, a review of existing game-based inoculation interventions against misinformation, begins to address the third research question. The final study of this thesis, the design and development of a new game-based intervention building resistance to digital social engineering attacks, complements findings from the second study in answering the third research question. In the below sections I introduce each study in turn, addressing how each study has been designed to answer its corresponding research question(s), and how the findings from each study informed the design of the next.

3.1 Exploring the Learning Potential of a Game-based Intervention to Inoculate Against Misinformation

This section introduces the methodology for the first study in the thesis. In this section I discuss how the study has been designed to answer research question 1 and 2, justify methodological design choices that have broken with existing game-based inoculation literature, and overview the limitations of the study. There are two discrete distinct elements to this study, a quantitative element and a qualitative element. Each element

addresses its own research question, and each have been presented in their own chapter in this thesis. Nonetheless, the quantitative and qualitative elements of this study are closely linked, and the qualitative findings have in many cases helped explain quantitative results, and (to a lesser extent) vice versa. As well as introducing the study intervention, this section individually presents the methodology for these elements. The present study is designed to answer the following research questions: how the effectiveness of analogue game-based interventions against misinformation changes over time, and what kinds of learning are stimulated within analogue game-based inoculation interventions against misinformation.

First, I considered whether, given existing literature in analogue game-based interventions, it would be more appropriate to use an existing analogue intervention or create a new one for this study. Recent replication studies have raised concerns that some game-based interventions may reduce belief in true as well as false content (T. Graham & Andrejevic, 2024; Modirrousta-Galian & Higham, 2023) or have failed replication entirely (Wong & Wu, 2023). As far as the author is aware, analogue game-based interventions have seen no replication at time of writing. Through replication, the present study could not only give a more contextualised contribution to literature but could draw on previous research to interpret findings. The collaborative card game *Fake News* (Roozenbeek & van der Linden, 2018) was chosen as an early and accessible example of an analogue game-based intervention against misinformation. As is outlined in the sections below, this study includes methods from this original research, but includes a longitudinal element, and qualitative elements including semi-structured interviews and session observations. This study is a replication with extension (Brandt et al., 2014). The study was not pre-registered due to its exploratory nature. This study was approved by the University of East Anglia School of Psychology Research Ethics Subcommittee on 09, 22 (ETH2223-0090) and 05, 23 (ETH2223-1878). Participants gave written consent before taking part in the study, and again before starting interviews. Below, I give an overview of the intervention *Fake News*.

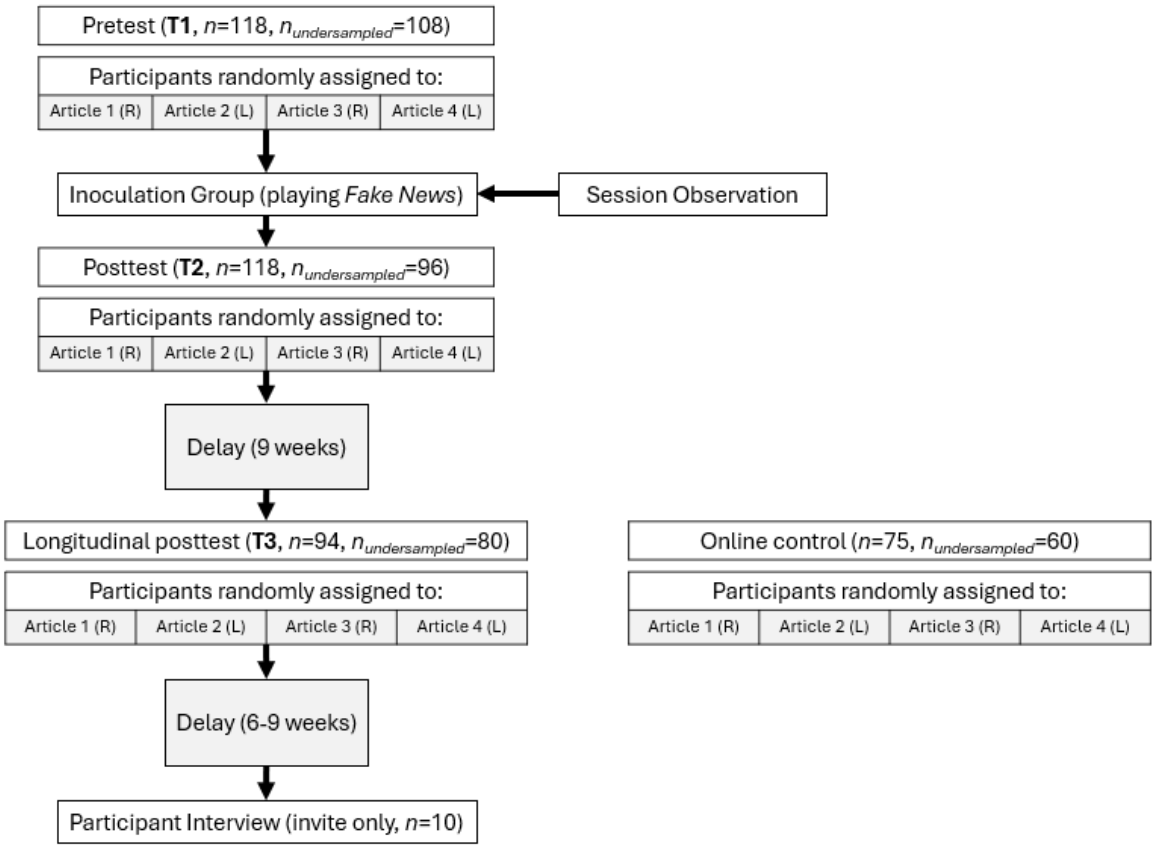


Figure 3: Study Design Flowchart

The Fake News Game

The intervention was a collaborative card game created by the University of Cambridge Social Decision-Making Lab and DROG called *Fake News*. Within the game, players work collaboratively to write a fake news article in the style of a given character by selecting and ordering pre-written article sections (see Figure 5). The game is set within a fictitious world which players are introduced to during the rules briefing, although the game focusses on the real topic of immigration. The game is designed to place "news consumers in the shoes of (fake) news producers (…) to think proactively about how people might be misled in order to achieve a goal" (Roozenbeek & van der Linden, 2018, p. 572). More information on the intervention *Fake News* can be found in the original publication by Roozenbeek and van der Linden (2018).

The game consists of three types of cards: source cards, character cards, and article cards. At the beginning of each session, the group is allocated one of four possible character cards (Dennis in Denial, Mike the Money Changer, Dana and Fox: Co-Conspirators, and Paula for Political Posturing). Each character has different motivations and techniques for manipulating public opinion. For example, the motivation for Mike the Money Changer is "his bank account, not to his readers, journalistic ethics, or overhyped niceties like 'the truth'" (see Figure 4). Other characters deny problems (Dennis in Denial), seek to create conspiracy (Dana and Fox), or "score

political points" (see Appendix 6). The techniques employed by Mike the Money Changer include overexaggeration of issues, selective reporting, and drawing on readers emotions. Players play collaboratively as a single character for the whole session.

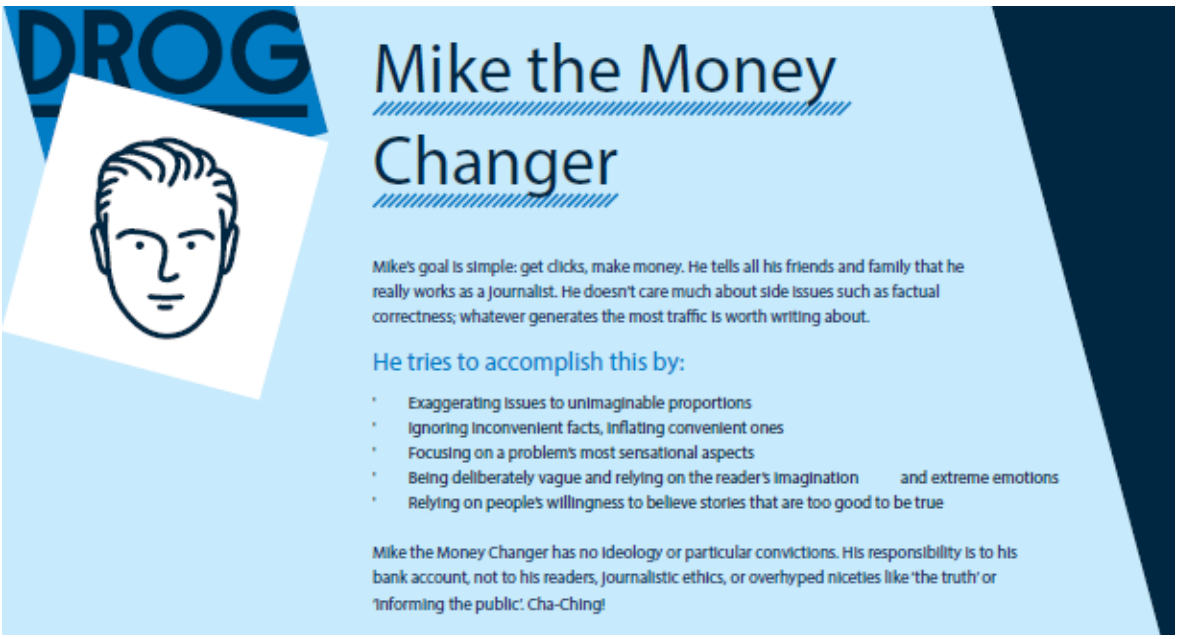


Figure 4: Character card for Mike the Money Changer, one of four possible character cards in the Fake News Game.

Article cards are the central elements to the intervention. Each card contains a pre-written article section that falls in one of ten ordered categories (order represents the order they will appear in the finished article): title, image, introduction, statistics, analysis of statistics, cause, underlying problem, expert opinion, and conclusion (see example card, Appendix 6). For each section there are four article cards that groups can chose from, with each of the article cards aligning with one of the four characters. Players work together to decide, for each category, which article card fits their given character. The content of four article cards for the 'statistics' card is shown in Table 2. A source card (the same for all player groups, regardless of character) introduces players to the true facts of the scenario. Although not a playable card, the scenario context is important for players, as understanding the true facts of the scenario gives important context when deciding which article card is more associated with their character. This source card (and other game cards) is provided in the appendix (see Appendix 6).

Table 2: Article card content for article category two, "The numbers". The character association of each card is not visible to the players during play. Players must collaboratively decide which article card correlates to their character.

Dennis in Denial	Mike the Money Changer	Dana and Fox: Co-Conspirators	Paula for Political Posturing
Incidents: excuses to fine immigrants In total, the CAA has registered 9.166 reports of incidents in detention centres across the country. While this constitutes a slight increase when compared to previous years, only 166 reports have led to actual prosecutions (1.8	A crisis is looming? The CAA's report is absolutely incredible. Despite all assertions to the contrary, the number of crimes committed by asylum seekers in detention centres has gone up by more than 300 percent. The consequences of this are looming large over our country.	What really happened? The CAA claims it has registered almost 10.000 cases of "incidents" involving asylum seekers in 2016, but declines to elaborate. The government now appears to be trying to keep the explanation behind the numbers a state secret.	An explosive rise There were 9.166 incidents involving asylum seekers recorded in 2016, against approximately 3000 in all of 2015, already a record. This led to an extreme burden on our police force, which had to be called in on 3.637 of these incidents.

Each groups' character is allocated randomly, which players keep for the entirety of play. The learning objectives of this intervention are ones of learning and repetition; players are informed of misinformation techniques central to their character and are then asked to repeatedly identify these techniques in written text when assembling their article. Once players had finished creating their article, it was compared to the correct article for their character by the session lead and a score is awarded (a maximum score of 9 is possible as there are only 9 article sections). A leaderboard can be created to track teams' scores. Groups can play in teams of 2-4, with gameplay lasting a maximum approximately 30 minutes. Figure 5 shows a typical tabletop layout at the end of a 30-minute session.

Figure 5: Collaborative analogue game Fake News, developed by Roozenbeek and van der Linden (2018) in collaboration with DROG.

Sample and Participants

The *Fake News* experiment was carried out at the University of East Anglia, in the East of the United Kingdom. Participants were recruited through a local research participation scheme, where members of the public can take part in (primarily psychology-based) university research for a small monetary incentive. Monetary incentives were used for both the pre- and post-test and the longitudinal post-test to reduce sampling bias (i.e., such that participants weren't just taking part due to an interest in misinformation) and to discourage attrition. Places to take part in the study were given on a first come first serve basis (convenience sample), and participants had to be 18 years or over and own a smartphone. During the signup booking process, participants submitted demographic information on age (using 10-year bands), gender, education, citizenship, and self-reported their political leaning (measured between 0 and 100, labelling 'Labour' on the left and 'Conservative' on the right). Given that the sample was self-selected, demographics were skewed from the general population. The age of participants was bimodally distributed, with most participants being either over 51 or between 18 and 24 (*Mdn* = 41-50). In total, 71% of the participants were female, and 29% were male. Most participants had achieved at least a college education (42%), and participants were on average more left leaning (μ = 34.6, σ = 28.6). Regrettably, there were few participants with nationalities outside of Western Europe (16.1%), populations that are in need of further investigation in this area (Harjani et al., 2023; Roozenbeek et al., 2024).

Quantitative Study Elements: Replication With Extension

The first element of this study has been designed to address the first research question: how does the effectiveness of an analogue game-based intervention against misinformation change over time. A traditional within-subjects design was chosen over between-subjects, as I felt that, given that participants were required to attend in person (and in some cases traveling from outside the city), it would be unfair that participants in a control group would not be able to take part in the intervention (N = 118). Consistent with existing longitudinal research with digital game-based interventions (Maertens et al., 2021), I implemented a 9-week longitudinal post-test (n = 94, 20.3% participant attrition). Participants completed the same short survey before (T1) and immediately after (T2) the game session to test the effectiveness of the intervention at inoculating against misinformation, and then again 9 weeks after the game session (T3) to measure changes in this effectiveness over time (see Figure 3). T1 and T2 were completed in person (supervised by the researcher), and T3 was completed remotely. To account for the potential of mode effects (the survey mode chosen can introduce unwanted bias, see 'mode effects', Heerwegh, 2009; Hochstim, 1967; Roberts, 2007), an online control test was also implemented (n = 75). Participants in the online control group read one

randomly assigned article and completed the same survey as within-subject participants. The data collection for this study was completed in two parts, initially in October 2022 (n = 54, N. Henderson et al., 2023), and then in August 2023 (n = 64). The online control was completed in May 2023 (n = 75). The study was not preregistered due to its exploratory nature.

T1, T2 and T3 (plus the online control) involved participants reading a previously unseen 'fake news' article and answering open-ended and 7-point Likert-scale questions. This follows a similar methodology previously used by Roozenbeek and van der Linden (2018) with the same intervention (following a replication with extension methodology). Participants were asked to read one of four randomly assigned[1] articles, which were written in a similar style, length, structure, and used the same number and type of common misinformation techniques (4 occurrences of hyperbole, 4 of 'common man' appeal, 2 arguments from authority, 2 uses of conspiratorial reasoning, one demonisation of the out-group, one use of whataboutism, and use of an ad hominem attack). All articles contained false arguments and incorrect information about immigration within the United Kingdom and the EU. All four testing articles are included in Appendix 7.

Despite only having three testing stages, four testing articles were written[2] to ensure an equal political balance (two were written as left-leaning and two were written as right-leaning, see supplementary materials). As well as following previous research (Roozenbeek & van der Linden, 2018) it was decided that participants should read *novel* 'fake news' articles in testing phases as, if using the same article each time, participants may be inclined to improve post-test scores after becoming more familiar with the terminology/topic (rather than measuring any improvement in misinformation resilience; for further discussion on this see Roozenbeek et al., 2021). Articles were randomly distributed to participants in testing phases but restricted such that participants would not read the same article more than once. After reading the 'fake news' article, participants were asked to answer 7-point Likert-scale questions on how persuasive they found the article, how reliable they found the article, how much they personally agreed with it, and how familiar they felt they were with the article's topic.

[1] The 'fake news' articles used in quantitative tests were randomly allocated to participants in such a way that they were equally distributed in each testing phase (for example, in the pre-test 30 participants read article 1, 30 read article 2, 30 read article 3, and 30 read article 4). However, due to some participants failing to attend and attrition in the longitudinal post-test, this distribution in our data is unbalanced. Where necessary, we have undersampled our data to balance this distribution by removing the most recent participant responses (this was chosen over random undersampling to enable replication of these results). Code used for this undersampling approach is provided in supplementary materials, and when referenced in this paper, we have clearly labelled the data as "most-recent undersampled".

[2] Two have been adapted from previous work (see Roozenbeek & van der Linden, 2018).

As well as following a replication with extension study design, meta-analysis evidence suggests that perceived effectiveness is a valuable methodology in circumstances where actual effectiveness is challenging to measure (Dillard, 1998). Participants were also asked to answer three open-ended questions: what was the writer trying to convince them of, which arguments they found persuasive/not persuasive, and which arguments they found reliable. Survey questions are included in Appendix 8.

To measure whether changes to the three main quantitative measures (persuasiveness, reliability, and how much the participant agrees with the article) was significant, Kruskal-Wallis tests were used. Although this breaks with convention (game-based inoculation research has frequently used ANOVA (or variations such as MANOVA and ANCOVA) tests to measure longitudinal change (see Appel et al., 2025; Maertens et al., 2021; Roozenbeek & van der Linden, 2019), this data was non-parametric in nature (data included responses on the four different 'fake news' articles) and thus a non-parametric test was used. Moderation and mediation analysis were also important analyses completed in this study. Moderation analysis (i.e., the measurement of how a variable affects, or moderates, the relationship between the independent variable and the dependent variable) allowed me to identify how key inoculation variables such as threat and issue familiarity affected the intervention's effectiveness to improve misinformation resilience. Mediation analysis (i.e., the measurement of how a variable acts as a mediary between the independent variable and the dependent variable) allowed me to analyse how perceived reliability mediated the effect of the intervention on perceived persuasiveness. Further moderation analyses completed not discussed in this thesis have been provided in the appendix (Appendix 3; Appendix 5). A bivariate Pearson correlation matrix with quantitative measures and demographic variables is also supplied in the appendix (Appendix 1). T-Tests were also used to measure whether the differences in measures between the follow-up post-test (T3) and the online control group were significant. Kruskal-Wallis and T-Tests were completed in Python (version 3.12.4) using SciPy[3] (version 1.13.1) and statsmodels[4] (version 0.14.5) modules. Moderation and mediation analyses were also executed in Python, and these findings were verified with SPSS (version 29.0) using the Processes macro (v5.0) as well.

Qualitative Methods

The second element of this study has been designed to address the second research question: what kinds of learning are occurring within analogue game-based inoculation interventions against misinformation. Much existing game-based inoculation research has relied on quantitative measures to assess intervention efficacy. Following a pragmatist research paradigm has led the design of the second element of this study to

[3] https://scipy.org/
[4] https://www.statsmodels.org/

break with this convention. This study introduced qualitative methods, including participant observation of game sessions which were recorded in a field diary for each group taking part, and semi-structured interviews with a reduced sample of study participants.

Interviews for the study were conducted 6-9 weeks after the longitudinal follow up had been completed, 15-18 weeks after taking part in the in-person intervention. All participants from the second round of data collection were given an opt-out option to decline being interviewed. From those remaining, 17 participants were approached for interview, guided by a sampling strategy which aimed to capture diversity in age, nationality, political leaning, and level of engagement during their complete session. 10 participants consented to interview; Table 3 shows demographic information for participants that took part in interviews and includes metadata on the interviews themselves. All interviewees referenced in this thesis have been given pseudonyms to protect their anonymity. Interviews were conducted in a university café and lasted approximately 40-60 minutes. Interviews were individual and semi-structured in nature, covering the following topics: recollections and overall impressions of the study; learning occurring throughout the test, including outside of sessions; The dynamics of the sessions themselves, including the game and participant's teammates, and how participants felt about this; and participants' broader experiences and views on misinformation and news.

Table 3: Demographic and interview details on interviewed participants

	Pseudonym	Age	Gender	Nationality	Political leaning	Interview date	Interview length
1	Alexander	18-20	Male	British	48	18/10/23	36 min
2	Amber	31-40	Female	British	37	17/10/23	45 min
3	Andrea	18-20	Female	Vietnamese	4	10/10/23	37 min
4	Brooke	Over 60	Female	British-American	17	10/10/23	51 min
5	Curtis	Over 60	Male	British	0	13/10/23	1 hr 0 min
6	Dominic	Over 60	Male	British	78	10/10/23	53 min
7	Julia	Over 60	Female	British	100	31/10/23	36 min
8	Robert	Over 60	Female	British	1	11/10/23	1hr 4 min
9	Joan	Over 60	Male	British	7	11/10/23	1hr 4 min
10	Wesley	Over 60	Male	British	0	27/10/23	56 min

For each interview the interviewer was accompanied by a number of documents and materials to assist with the interview processes. First, session observations made during the session were printed and reviewed by the interviewer beforehand to ensure the interviewer was informed when discussing the participant's experiences of the intervention and interacting with other players. Other materials

were brought to the interview for reference during the interview (where needed), including cards from the game-based intervention *Fake News* itself, and the interviewees responses to survey questions at all three testing stages (T1, T2, and T3). As the interview was semi-structured in nature, the interviewer was also assisted by an interview guide. Following guidance from Bryman (2012), the interview guide included a number of direct questions, follow-up questions, and probing questions. Example questions from the interview guide are provided in Table 4, although the full interview guide is provided in Appendix 9. As well as being an important material for semi-structured interviews generally, the interview guide was a helpful material to assist me as a neurodiverse researcher.

Table 4: Extract from guide for semi-structured interviews of participants having previously participated in Fake News sessions.

Introducing questions	• You played the game just over 9 weeks ago, what do you remember from the session? • Why did you decide to take part in the study? • What were your overall impressions from the session?
Direct questions	• How do you feel your biases are influenced by your news consumption? • How did playing the game affect you and your day-to-day actions, thoughts, and processes? • Would you play this game again? • How did you feel the number of players affected the session?
Follow-up questions	• Get the interviewee to expand on their previous answers. Why? How? In what way?
Probing questions	• Follow up through direct questioning. For example, "could you expand on this?", or "you said [this], what do you mean?"
Specifying questions	• Very similar to probing questions but specifying detail. For example, "what did you do then?".

The qualitative coding approach taken was inductive in nature, extracting themes as they were revealed from the data. Interviews were transcribed verbatim using Microsoft Word Dictate (with human intervention), and qualitative coding was completed using NVivo (version 14.23). During the interview processes reoccurring themes had already emerged, such as the positive sentiment around analogue intervention design, and many participants' personal experiences with misinformation. After the 10 interviews had been transcribed these provisional themes informed the coding processes, although I sought to remain mindful that this may prejudice the coding processes. I first undertook an open coding process (Anselm & Corbin, 1990) where, through careful examination of transcripts, seven descriptive, conceptual codes emerged: session recall, reason for participation, evaluated value, general game design impressions and proposed improvements, group interactions, news consumption practices, and methodology meta-discussion. Throughout this coding processes, analytic memos were written as research themes began to emerge which tied findings from open coding to the study's literature, including themes on participant learning,

participants' prior knowledge of misinformation, the role of inoculation theory, and intervention design assumptions and evaluation. An excerpt from a coded transcript is provided in Appendix 10 to illustrate this qualitative data analysis processes. Contextualised quotes from these interview transcripts are used to develop and illustrate arguments in Chapter 5: they appear in quotation marks with the participants' pseudonym (see Table 3) in brackets.

During the second round of data collection, participant observations were carried out during game sessions to better understand how groups interacted with intervention materials, each other, and how groups worked collaboratively to complete the intervention. Initial observations completed during the session were in 'jotted note' form (Bryman, 2012; Lofland & Lofland, 1995): brief notes and comments on events and behaviours recorded in note form. After each session had finished, these observations were written up as full field notes, including my own reflections and immediate analysis on the session. Full field notes for each session varied in length from 73 to 267 words depending on notable observations, behaviours, analysis, and time available (before the next session started). These notes were used to support the sensemaking stage of qualitative analysis of interviews, and to contextualise interview quotes and vignettes. A table of all the study's sessions (with accompanying meta-data) from which observational field notes are drawn can be seen in Appendix 12.

Limitations

Despite some promising findings, this study suffered from some methodological limitations. Firstly, the sample used in the study is a convenience sample, and not representative of any given population. Qualitative interview findings from the study demonstrated that many participants had taken part in the study out of personal interest in misinformation and analogue games. Random sampling is especially difficult in in-person research, and it is important to highlight this for this study. Second, semi-structured interviews were completed between 16-18 weeks after the intervention. This methodological design aimed to understand the longitudinal recall of participants over time but may have negatively impacted the accuracy of responses and discussion. There were times where participants mis-remembered elements of the intervention and had to be corrected. Conducting interviews closer to the intervention may have allowed a deeper insight into participants' in-session experiences. Perhaps varying the time between the intervention and the interview, as seen in the longitudinal quantitative research by Maertens et al. (2021) could have given a more comprehensive longitudinal view of the intervention's effects. Thirdly, session observations and participant interviews were only completed during the second round of data collection (n = 64), meaning that session observations completed were not necessarily representative of the whole sample. Fourth, some aspects of the quantitative analysis

could have benefited from greater statistical power, particularly when exploring unbalanced demographic variables (e.g., nationality, political leaning, and gender). The study was hampered by a high non-attendance rate; given the in-person nature of the study, and that the game *Fake News* could be played by a maximum of four players, participants had to book slots in which they were available and willing to travel into the university. The non-attendance rate for this part of the study (excluding follow-up interviews) was 33.7%, with an attrition rate of 20.3% for the longitudinal post-test. This had a great effect on the power of the study, limiting the number of causal conclusions that could be drawn. Better anticipation of participant non-attendance and attrition may have improved the power of statistical tests completed. Finally, although a socio-demographically balanced cross-section of participants were reached out to for interview, the participants who accepted were older, and more left leaning ideologically than the overall sample. Nonetheless, I believe that the contributions remain valid despite these limitations, and the limitations of the sample were taken into account when interpreting the qualitative data.

This section has introduced the methodological design process for the first study of the thesis, a mixed-methods, within-subjects, longitudinal study with the game-based intervention *Fake News*. This section has introduced the intervention itself, as well as both quantitative and qualitative methods employed to measure the intervention's immediate and long-lasting effectiveness. Furthermore, this section has introduced methods used to analyse session observations and interviews, as well as the statistical methods used to analyse quantitative data. Finally, this section has discussed the limitations of the study, and the potential impact these limitations had on the study's findings. Materials used in the study (e.g., intervention pieces, survey questions, 'fake news' testing articles) can be found in the appendix of this thesis.

3.2 Inoculation Theory as a Design Approach to Game-based Misinformation Interventions

This section introduces the methodology for the second study in the thesis, a review of game-based inoculation interventions against misinformation. In this section I discuss and justify how the study has been designed in part to answer research question 3, how its design has been informed from results of the first study (following a consecutive research strategy) and overview the limitations of the review. This study reviewed game-based interventions against misinformation, exploring both the game design trends of the interventions, and how the interventions drew on inoculation theory literature. Importantly, this review not only assessed the serious games themselves (through playtesting, where available), but of the interventions' corresponding publications as well.

In answering research questions on longitudinal effectiveness and types of learning the first study revealed some unexpected findings. Notably, some aspects of intervention design (i.e., analogue, collaborative design) played a leading role in the outcomes of the intervention. Moreover, qualitative findings from the study questioned how inoculation had been drawn on to inform the design of the intervention; the intervention did not appear to elicit threat in players, and interviews revealed that many participants were already aware of misinformation techniques used[5]. These findings have brought into question a major assumption for game-based interventions, that their theory-driven design draws on inoculation theory literature. The third research question of this thesis asks how game-based interventions may be better designed for more effective outcomes. Following a consecutive research strategy, it is clear that to improve inoculation-driven game-based design processes, a deeper understanding of existing design decisions and processes is required. By reviewing the design decisions and processes of existing game-based interventions, trends, inconsistencies, and research gaps can be identified, informing design processes for future practitioners.

Narrative reviews have been criticised for being less informative and reliable than systematic or meta-analysis reviews. These types of review, however, target specific, narrowly defined research questions. The research question this review seeks to address is broad and demands an inductive research approach. Therefore, an integrative narrative review was chosen. Integrative narrative reviews allow for the inclusion of broad types of research (e.g., experimental, non-experimental, empirical, and theoretical). This is essential as this review includes not only analysis of publications but involves the playtesting and analysis of game-based interventions themselves to allow a more comprehensive understanding of this specific type of gamified misinformation intervention. Integrative reviews are also completed for a wide range of purposes; not only to review evidence, but methodological approaches, theories, and more. Whittemore and Knafl (2005) define the processes of an integrative narrative review to be completed in five stages: problem identification, literature search, data evaluation, data analysis, and presentation. This review has sought to follow these steps to maintain transparency and unbiased.

This study reviewed and analysed 25 game-based inoculation interventions that aim to build resistance to misinformation. Interventions and their corresponding publications were identified through a snowball identification method. First, the chosen

[5] Although some of these findings may have been due to a convenience sample and not entirely reflective of the *design* of the intervention (as the intervention had been effective in previous research with adolescents), I believe these findings still prompt a closer inspection of inoculation in game-based interventions.

database (Google Scholar was chosen due to its broad reach) was searched with terms that aligned to the review's aim, including "game-based", "inoculation", and "misinformation". Associated synonyms were also added to searches. Next, a snowball identification processes took place; eligible publications that presented game-based interventions against misinformation were analysed to see if they referenced other relevant literature. It is through this snowball identification processes that many interventions and publications were identified, as it was discovered that reference to inoculation theory was often superficial and at times not included in the title or abstract of the publication (see Chapter 6 for further discussion on this).

I have restricted this review's scope to a) games that have at least one corresponding publication, and b) games whose associated publications explicitly describe the game as being an inoculation intervention. Although game-based interventions such as *Facticious* (Grace & Hone, 2019) and *Fakey* (Micallef et al., 2021), may inadvertently stimulate core inoculation mechanisms, they are not presented as inoculation interventions in their corresponding publications and have therefore been excluded from this review. This is to direct discussion to the interpretation of inoculation theory in game-based contexts (for wider reviews of misinformation games, see DeJong, 2023; Kiili et al., 2024). Online preprints of yet-to-be published work have been included in this review, and games in languages other than English are included if the corresponding publication is written in English. In cases where multiple editions of the same game exist (e.g. *Cranky Uncle Vaccine*, *Radicalise*, and *Bad Vaxx*) these have been listed separately.

Following integrative review methodology outlined by Whittemore and Knafl (2005), the variables with which the game-based interventions and their corresponding publications are reviewed against were defined before data collection began[6]. The variables through which game-based interventions and their corresponding publications are reviewed can be placed into two categories. Firstly, the design of the game-based interventions must be reviewed to reveal trends and areas of under exploration in intervention design. To this end, this review assesses interventions' topic (i.e., whether the intervention seeks to inoculate against a specific topic), genre (i.e., game design characteristics), platform (e.g., analogue, browser, or mobile), target audience, and roleplaying perspective (i.e., whether the player plays as a misinformation creator or 'defender'). These variables have been chosen as some have received recent attention from scholars seeking to understand their role in intervention outcomes (see research on roleplaying perspective, Appel et al., 2025). Secondly, given findings from the first study questioning the role of inoculation on intervention

[6] I note that following a pragmatist research paradigm (and the research being exploratory in nature) allows for some flexibility in these variables where appropriate.

outcomes, this review also seeks to better understand how inoculation theory has been drawn on in the design processes for existing game-based interventions. These variables address both the type of inoculation intervention, and how inoculation-driven design has been discussed in the corresponding publication. Variables chosen are threat source (i.e., implicitly through materials or explicitly through a forewarning), counterarguments type (i.e., issue-based or technique-based), the existence of other related components (e.g., debriefing components, see Leder et al., 2024), inoculation design considerations, and game design processes. The review turns more to corresponding publications as well as game-based interventions to assess this second category of variable.

Playtesting of browser games took place between the 23rd of September and 13th of December 2024 in the Chrome browser, although this was not possible for games that were not publicly available. Analogue games (as well as *Unnamed 3*) were not playtested[7]. Unfortunately, only 14 of the 25 reviewed game-based interventions were playtested; four of those not playtested were analogue, and seven of the remaining digital game-based interventions were not publicly available. Importantly, some interventions that were publicly available at the time of playtesting are no longer publicly available at time of writing. Where interventions were not publicly available (and I was therefore not able to playtest), data evaluation was completed from corresponding publications. The playtesting processes involved the researcher playing each game-based intervention to completion at least once, until the researcher was satisfied that all possible data for all variables had been collected. The full dataset of all games reviewed, including variables not discussed in this review, can be found in the supplementary materials at https://doi.org/10.6084/m9.figshare.29310692.

This section has introduced the research methodology for the second study of the thesis, an integrative narrative review of game-based inoculation interventions against misinformation. Specifically, this section discusses how findings from the first study of the thesis (particularly around the role of inoculation and learning) informed the design of this review. Furthermore, justification for the review methodology, literature search terms and process, inclusion and exclusion criteria, and how each game-based intervention (and corresponding publication) was reviewed has been outlined and discussed through the lens of a pragmatist inductive research paradigm.

[7] The analogue intervention used in the first study of this thesis *Fake News* is included in the review and marks the only exception to this.

3.3 *The disPHISHinformation Game*: Creating a Serious Game to Fight Phishing Using Blended Design Approaches

This section introduces the methodological approach taken in the third study of the thesis, an exploratory study designing and developing the game-based cyber security intervention *The disPHISHinformation Game*. In this section I discuss how the motivation for this section to contribute to the third research question was driven by findings from the previous review, justify the methodological design choices of the study, and highlight the study's limitations. This study sought to create a game-based intervention to inoculate players against social engineering cyber security attacks such as email and SMS phishing. Contribution of this study comes not only from the intervention itself but the process through which it has been designed and developed, both of which have been heavily influenced by the findings of the review outlined in the previous section. Data collection for this study was given on 15th May 2024 by the UEA Faculty of Science Research Ethics Subcommittee (ETH2324-1332).

This study sought to join the previous section in contributing to the third research question: how could game-based inoculation interventions against misinformation be better designed. The second study of this thesis, an integrative narrative review of game-based inoculation interventions against misinformation (see previous section), found that many design genres and platforms, such as interventions that were in-person (analogue) or used virtual reality, was under explored. The review also found a lack of transparency surrounding design decisions and processes, and that many game-based interventions drew superficially on inoculation theory literature to inform their design. In addressing the research question, the review found that game-based inoculation interventions hardly interacted with inoculation theory mechanisms at all. This third study of this thesis does not outline a prescriptive process which game-based inoculation scholars should follow. Instead, this exploratory pilot study contributes to the review's findings by demonstrating how game-based inoculation interventions may be better designed than existing literature.

Digital social engineering (i.e., malicious messaging often received over email or SMS intended to deceive the recipient) was chosen as the topic for the intervention as a) it is a pressing issue responsible for loss[8] to individuals and businesses (*Threat hunting report*, 2025; Rizvi & Fordham, 2025), b) it shares many similarities with misinformation (e.g., manipulative techniques used in attack messages), and importantly c) it presents an opportunity for game-based inoculation that may not

[8] Although the impact of phishing attacks is often stated as being largely financial, falling for some attacks can come at great personal cost too (e.g., see Z. Wood, 2025).

suffer from inoculation hesitancy in the same way that game-based misinformation interventions do. It has become increasingly common for organisations (particularly with internet-facing products or systems) to offer security training for employees; however, the risks have remained high (Coutinho et al., 2023; Rizvi & Fordham, 2025), and some have called into question the effectiveness of conventional phishing training (Fernando & Arachchilage, 2019). This study sought to meet this need through organisational engagement with AVIVA, a large multinational insurance company with offices located in Norwich, UK. Adopting engaged research practices allows scholars to interact with partners and their diverse perspectives to inform more equitable and effective research to meet societal challenges (Bednarek et al., 2025; Cunliffe & Scaratti, 2017; Holliman, 2017). AVIVA, like many multinational large enterprises, faces a high volume and wide spectrum of different digital social engineering attacks, and their knowledge and organisational experience helped inform elements of the intervention's design.

The contribution of this study comes not only from the intervention itself, but the processes through which the intervention was designed. Following findings from the second study of the thesis (an integrative narrative review of current game-based interventions against misinformation), I sought to adopt a more transparent design process. This meant not only a more inoculation theory-led design, but working with an established serious game design framework in the design phase of the intervention and following recognised game development processes to realise this design. Serious game design is a highly multi-disciplinary role (Culyba, 2018), and as the researcher for this PhD project, introspection was important for me to reflect and acknowledge my own knowledge and training as a game-based intervention designer. As outlined by Jesse Schell "[game design] is a delicate art that must be practiced — and with practice your skill at it will improve" (Schell, 2019, p. 18). To help inform and guide the intervention design processes, this study draws on The Transformational Framework (Culyba, 2018). Although other serious game design frameworks have been developed (Annetta, 2010; Ávila-Pesántez et al., 2017; Kiili et al., 2012; Miller et al., 2024; Robson et al., 2015), these offer less instructive assistance to serious game developers.

Through adopting an iterative development approach (following 'playcentric' development principles, see Fullerton, 2004), this study included an in-person public playtest. This mixed-methods engaged research approach sought to not consider people as subjects, but as partners in the design processes (a key consideration in iterative game design processes, see Schell, 2019). This methodological design of this playtest aimed to both measure the perceived (self-reported) level of learning facilitated by the intervention, and record player feedback on intervention design that could help to further inform the intervention's design. Data collection was completed

at AVIVA's Norwich offices (N = 13) and was completed in two rounds (n = 7 and n = 6). Before the session began, participants entered information on age and highest level of education achieved. Both of these socio-demographic variables have been shown to affect how individuals classify messages (e.g., classification speed, level of caution, and in some cases accuracy, see Gavett et al., 2017; Sarno et al., 2020; Sarno & Black, 2024), and although not used in quantitative analysis in this study (due to the small sample size) they provide context to qualitative insights given. Further contextual information was also collected on frequency of computer use, and previous experience playing games.

Next, participants were asked how cyber-aware they felt, and what features indicate that a message may be phishing (open-ended question). Cyber awareness was measured on a 5-point Likert scale, although points on the scale were described within the context of the intervention (1 = *Absolutely not cyber-aware, little understanding of phishing*, 2 = *Not very cyber aware, have very occasionally thought about phishing*, 3 = *Somewhat cyber-aware, have sometimes felt unsure about phishing*, 4 = *Cyber-aware, usually on the lookout for phishing*, and 5 = *Very cyber-aware, always on the lookout for phishing*). These two questions were asked both before and after participants completed the intervention to measure a) the level of alertness, awareness, or threat[9] of phishing attack felt by the player, and b) the different types of learning (e.g., factual, semiotic, procedural; see types of transformations, Culyba, 2018) experienced by players. After the session players were also asked to complete open-ended questions on their experience taking part in the intervention: how clear were the instructions provided with the game; what elements did you enjoy; what would you like to see more of in the game; do you feel that there was anything missing from *The disPHISHinformation Game*. Open-ended questions were chosen to give participants freedom of feedback around a number of design aspects of the intervention that may help identify areas of design that work well, require improvement, or could be drawn on more in the intervention's design. Participants' responses and the observation of playtesting sessions were analysed both categorically and inductively, identifying key themes and quotes which led to changes in the design of the intervention.

Although the empirical elements of this study use a very small sample size, this is a consequence of the methodology adopted. Building a working relationship with an external organisation enabled an informed design approach to meet an organisational need and offered a context through which intervention uptake was assured. Nonetheless, undertaking this study with an external organisation also limited the

[9] Although this measure of threat is not consistent with inoculation theory literature (e.g., Compton & Ivanov, 2012), this measure of threat is contextualised within the topic of the intervention.

study. Firstly, this research relied on the external organisation's willingness to engage, which is not guaranteed particularly as organisational incentives are often profit-driven. More iterations of public playtesting and prototyping would have better aligned with 'playcentric' design principles (Fullerton, 2004), but the organisation with which I was engaged was not always able to facilitate this. It is important to highlight that comprehensively adopting a 'playcentric' design approach while considering serious objectives is both long and costly (Culyba, 2018), and I have sought to adopt these methodologies given financial and time constraints of the project. Nonetheless, informed from earlier findings from this thesis, this study attempts to not only present a new game-based intervention, but demonstrate better informed design, development, and theory-led (i.e., inoculation theory) processes that may result in more effective interventions.

This section has introduced the research methodology for the third and final study of this thesis, the development of a game-based intervention inoculating players on digital social engineering cyber security attacks. Many of the contributions from this study come from the design and development methodologies adopted in creating the intervention. Thus, this section focuses instead on how findings from the previous study informed the third study's methodological design, the justification for using The Transformational Framework (Culyba, 2018) in the intervention's design, and both the importance and limitations of adopting an organisation-engaged research approach. Finally, this section introduces the methodological design for a public playtest with the intervention.

Summary

This chapter has outlined the research approach, and the various research designs and methodologies adopted in the research presented in this thesis. This thesis has adopted an inductive pragmatist research paradigm, employing a wide range of qualitative and quantitative methods to contribute to literature on inoculation theory and inoculation theory's application in game-based contexts. To address the three research questions of this thesis I have designed three studies: a within-subjects longitudinal study with the game-based analogue intervention *Fake News*, an integrative narrative review of existing game-based inoculation interventions and their corresponding publications, and the design and development of a new game-based intervention to inoculate against cyber security social engineering attacks, *The disPHISHinformation Game*. The findings from these studies have not only addressed research questions but, following a consecutive research strategy, have informed the design of successive studies in this thesis. This chapter has not only introduced the methodological design of each study but justified methodological design choices and highlighted potential limitations of each study.

In the following chapters of this thesis, I present and discuss findings from each study. First, the quantitative and qualitative elements of the first study of this thesis are presented in separate chapters (although quantitative results are discussed consecutively in light of qualitative findings). Chapter 6 presents and discusses findings study 2, and integrative narrative review of game-based inoculation interventions, and chapter 7 not only introduces a new game-based intervention but discusses the design decisions and processes adopted in its creation.

4. The Effectiveness of an Analogue Game-based Intervention Against Misinformation

Parts of this chapter are taken from the paper by the author of this thesis 'Exploring the Learning Potential of a Game-based Intervention to Inoculate Against Misinformation', soon to be submitted to Social Media + Society.

Quantitative studies investigating the effectiveness of game-based inoculation interventions have shown promising results (Kiili et al., 2024; Roozenbeek et al., 2024), but success is not guaranteed (Harjani et al., 2023), and very little is understood about how the resistance conferred from these interventions decays over time (Maertens et al., 2025). This chapter presents quantitative insights into the effectiveness of the analogue, game-based inoculation intervention *Fake News*, assessing both immediate and longitudinal efficacy. This chapter takes a quantitative, frequentist approach to evaluating intervention efficacy; Chapter 5 then looks more closely at insights provided by a qualitative methodology and calls on results from this Chapter in its consideration of the intervention. This chapter aims to answer the first research question:

> **RQ1.** How effective are analogue game-based interventions against misinformation, and to what extent does this change over time?

The findings presented and discussed in this chapter are drawn from the within-subjects study assessing the efficacy of the game-based inoculation intervention *Fake News* with adults in the East of England (completed June 2023). Individuals completed the intervention in groups of 2-4 in person, and their resilience to misinformation was assessed through a survey completed immediately before the intervention, immediately after, and again after 9 weeks. At the start of each survey participants read a 'fake news' article on the topic of immigration and then rated their perceived persuasiveness and reliability of the article, and how much they agreed with the article, on a 7-point Likert scale. Additionally, individuals reported their familiarity with the issue and answered open-ended questions on arguments they found to be reliable and persuasive. To reduce mode bias (longitudinal post-test surveys were

completed online, whereas the pre-test and immediate post-test was completed in person) an online control group was also undertaken (completed May 2023), the results of which are additionally assessed in this chapter.

The quantitative results from the study presented in this chapter failed to replicate original findings by Roozenbeek and van der Linden (2018). This chapter argues that the null effects observed from the three main Likert-scale measures may be a) due to the high participant issue familiarity (i.e., prior knowledge and awareness of the intervention topic, immigration), b) floor effects in quantitative measures, c) a lack of threat elicited by the intervention, and d) that the outcomes experiences by participants may have been broader than the quantitative measures used in this study. 36% of individuals reported being at least "somewhat familiar" of the article topic, and individuals had a significantly lower baseline persuasiveness measure when compared to the control group of original research with adolescents in the Netherlands (Roozenbeek & van der Linden, 2018). Other important findings relate to socio-demographic variables with the intervention descriptively reducing in-group agreement and out-group animosity with 'fake news' articles.

The four sections of this chapter each discuss the effectiveness of the game-based intervention *Fake News* through different lenses. First, the results from the intervention are considered in the context of inoculation theory, looking not only at base measures, but important moderating variables in the inoculation literature such as issue familiarity and affect (Banas & Rains, 2010; Compton, 2012). Following this, the chapter considers the present study as a replication of Roozenbeek and van der Linden (2018), making a direct comparison between present results with UK adults versus previous research with Dutch adolescents (Roozenbeek & van der Linden, 2018). The third section of this chapter more deeply investigates socio-demographic variables, assessing not only how these variables affected the efficacy of the intervention, but how socio-demographic variables such as nationality, age, and political ideology influenced how individuals' perceived the persuasiveness, reliability, and their agreement with 'fake news' articles. The final section of this chapter considers the *Fake News* intervention as a 'white box' system, going beyond intervention outcomes to consider how groups' score, level of interaction, group size, and the character played as impacted the conferred resistance from the intervention.

4.1 *Fake News*'s Efficacy as an Inoculation Intervention

Game-based inoculation research frequently focusses only on base measures to determine efficacy (Kiili et al., 2024; Roozenbeek et al., 2024), overlooking important moderators that are central to the inoculation theory (Compton, 2012; Pfau et al., 2005). Investigating these moderating variables (i.e., issue familiarity and affect) can provide

not only a clearer picture when evaluating the efficacy of this intervention but can contribute to the literature on inoculation theory itself (where these variables have been highlighted as promising avenues for further research, see Compton, 2024). In this section I first assess the within-group differences between results for the three Likert-scale measures. Next, the online control group is investigated in relation to the longitudinal post-test. Following this, I consider issue familiarity within the context of inoculation theory and investigate its role as a moderating variable. Finally, a sentiment analysis is completed on responses to open-ended questions to evaluate individuals' affective responses to the intervention. The section concludes with a discussion and interpretation of these results, leading to the next section comparing these findings to previous research with the same intervention.

The quantitative results of the study do not indicate that participants became more resilient to misinformation after completing the intervention. Previous longitudinal research with game-based interventions have demonstrated an immediate post-test improvement with some erosion at the longitudinal post-test stage (Maertens et al., 2021). There was a descriptive reduction of mean scores for persuasiveness judgements of testing articles between the pre-test (T1) and immediate post-test (T2), suggesting an improvement in participants' abilities to identify fake news techniques. However, mean scores for reliability judgements and personal agreement with fake news articles, in fact, increased, which undermines this apparent trend. Table 5 shows mean and standard deviation measures for all testing phases, between judgement types. A Kruskal-Wallis[10] test was completed on persuasiveness ($H(2) = 1.25$, $p = 0.5$), agreement ($H(2) = 0.52$, $p = 0.8$), and reliability ($H(2) = 1.09$, $p = 0.58$) judgements and found all quantitative changes to be non-significant (see Appendix 2).

Table 5: Mean participant persuasiveness, agreement, and reliability judgements of testing articles, with standard deviation in parentheses. To ensure an equal distribution of articles per test, data is 'most-recent' undersampled.

	Participant article judgements		
Time	**Persuasiveness**	**Agreement**	**Reliability**
Pre-test (n = 108)	3.43 (1.99)	3.36 (1.93)	2.81 (1.56)
Post-test (n = 96)	3.35 (1.75)	3.54 (1.98)	3.04 (1.61)
Longitudinal post-test (n = 80)	3.61 (1.57)	3.45 (1.80)	2.94 (1.61)

Figure 6 shows these measures for T1, T2, and T3. Graph A illustrates the perceived persuasive of 'fake news' articles between the three testing stages. Perceived

[10] Quantitative results from the Likert measures used judgements from four different articles of differing political leaning and were therefore non-parametric when combined. Using non-parametric tests for significance is in contrast to common practice within related literature, but I feel it is more robust with data of this nature. Additionally, quantitative data was 'most-recent' undersampled such that the distribution of testing articles in each testing phase was equal.

persuasiveness reduces immediately after the completing the intervention (T2), but this change is not significant. Graph B (Figure 6) shows how participants' agreement with 'fake news' articles shifted during the study. During the course of the study participants agreed with 'fake news' articles more, undermining the expected direction for this measure, although this change was also not significant. The final graph in Figure 6 shows how perceived reliability of 'fake news' articles changed during the study. This too shifted in an unexpected direction, increasing after the intervention (although this too was not significant). Participants also perceived 'fake news' articles to be less reliable than persuasive or agreeable, a trend observed by Roozenbeek and van der Linden (2018). This may be a symptom of the 'fake news' articles being on immigration; although a participant may agree with the article and find it persuasive, they may rate its reliability lower considering the polarising nature of the topic. Notable too are the proportion of participant responses that select the maximum or (more frequently) minimum response options. The suggests the existence of floor effects in these measures, which may be partly responsible for these null effects.

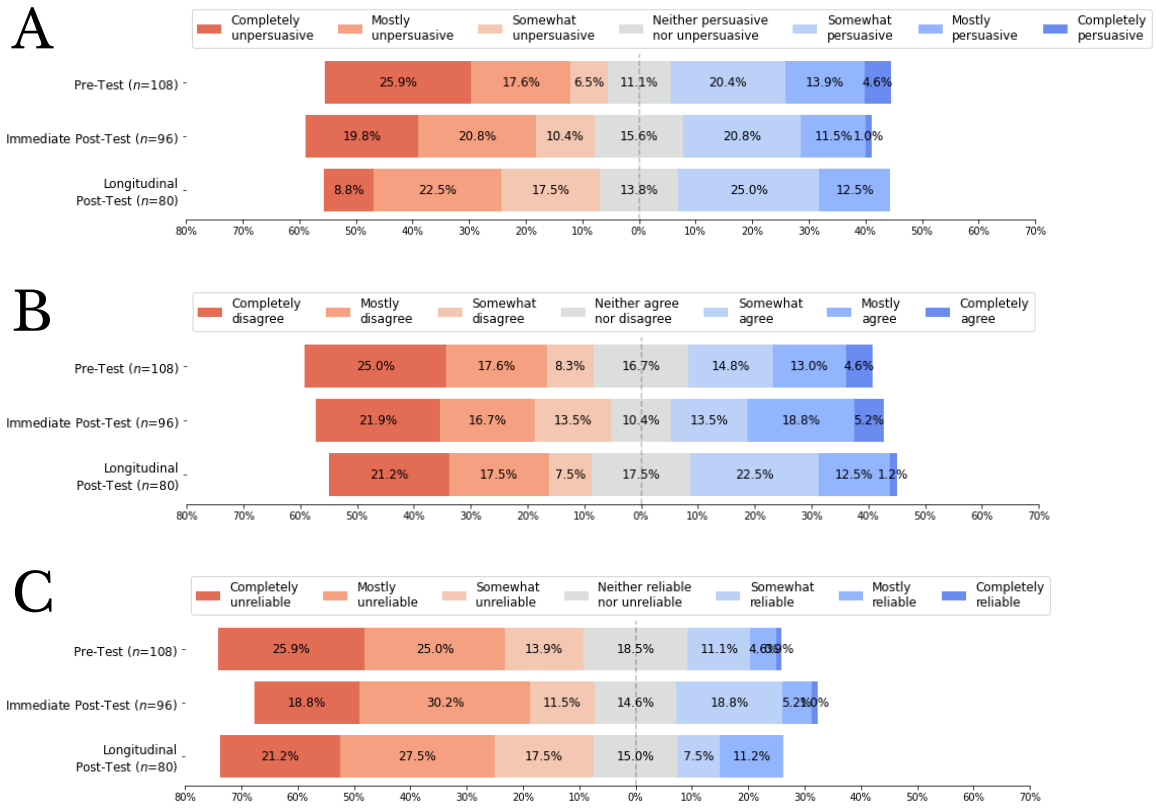


Figure 6: Participant persuasiveness, agreement, and reliability judgements of 'fake news' articles across all three testing stages. Data is given as percentages to scale the longitudinal post-test (after longitudinal participant attrition). To ensure an equal distribution of articles per test, data is 'most-recent' undersampled.

Although the post-test follow up was identical to the pre-test and immediate post-test in question, instruction, and response measures (note that, as discussed in Chapter 3, articles were randomly assigned in testing stages), existing research into quantitative methodology and survey modes advises that the mode of the survey (i.e.,

whether it was completed in person or online) can introduce unwanted bias (also known as 'mode effects', see Hochstim, 1967; Roberts, 2007), particularly when dealing with sensitive data (Burkill et al., 2016). Where the immediate pre- and post-test surveys (T1 and T2) were completed in person supervised by the researcher, the longitudinal post-test (T3) was completed by participants online at home. An online control group was completed to control for this change in survey mode, allowing a robust comparison to the longitudinal post-test (see Figure 7). Table 6 shows the means and standard deviations from all three measures, between these two groups. A one-tailed Independent T-Test showed no significant difference between the longitudinal post-test and the online control group for any of the three measures ($t(138) = -0.73$, $p = .23$, $t(138) = 0.32$, $p = .62$ $t(138) = -0.65$, $p = .26$ for persuasiveness, agreement, and reliability measures respectively). These null results of both within-subject differences and to the online control are unexpected as a) game-based inoculation interventions have been found to be effective in other research, and b) they differ to previous research completed with this game-based intervention (i.e., the card game *Fake News*) in the Netherlands (Roozenbeek & van der Linden, 2018). A closer look at moderating variables is necessary to better understand these null quantitative results.

Table 6: Mean participant persuasiveness, agreement, and reliability judgements of testing articles, with standard deviation in parentheses. To ensure an equal distribution of articles per test, data is 'most-recent' undersampled.

	Participant article judgements		
Time	**Persuasiveness**	**Agreement**	**Reliability**
Longitudinal post-test (n = 80)	3.61 (1.57)	3.45 (1.80)	2.94 (1.61)
Online control (n = 60)	3.82 (1.71)	3.35 (1.89)	3.11 (1.57)

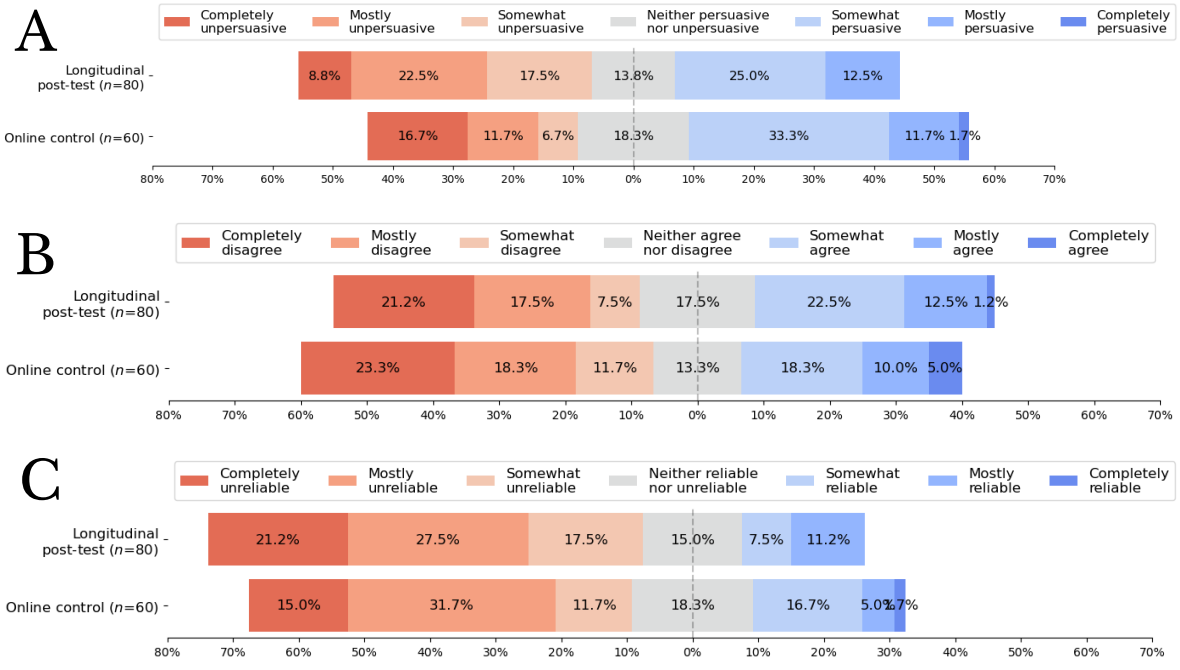


Figure 7: Participant persuasiveness, agreement, and reliability judgements of 'fake news' articles between longitudinal post-test and online control groups. Data is given as percentages to scale the longitudinal post-

test (after longitudinal participant attrition). To ensure an equal distribution of articles per test, data is 'most-recent' under sampled.

The inclusion of the issue familiarity measure in this study serves a dual purpose: to confirm balance between treatment groups and understand how issue familiarity serves as a moderator in this intervention. The study itself focussed on immigration in the United Kingdom; both the game-based intervention and the 'fake news' articles used this topical issue. Early inoculation literature (Pfau et al., 1997, 2005) suggested issue involvement played a key role in whether an intervention can confer resistance to an individual: too little, and the individual is unlikely to perceive that their attitudes are vulnerable to attack (or not care); too high, and the individual is already fully aware that their attitude is susceptible to attack (see for a wider discussion, Banas & Rains, 2010). A meta-analysis of inoculation theory by Banas and Rains (2010) found no significant relationship between issue involvement and resistance, but findings did trend in the expected direction, and subsequent reviews have continued to highlight its importance as a moderator (Compton, 2012).

Although those with a low issue familiarity were most affected by the intervention, this was not significant due to the low sample size at this issue familiarity level. In fact, only 14% of participants stated they were "mostly" or "completely unfamiliar" with the topic. To test the effect that issue familiarity has on the relationship between completing the intervention and perceived persuasiveness of 'fake news' articles, a moderation analysis was conducted (R^2 = .02, F(1, 237) = 11.15, p < .05). The moderation analysis found no significant effect between the intervention group and perceived persuasiveness, although perceived persuasiveness was descriptively lower after completing the intervention (β = 0.15, p = .27). The moderation analysis also found that individuals with a higher issue familiarity had a lower perceived persuasiveness, although this was also not significant (β = -0.12, p = .23). The interaction between the intervention group and perceived persuasiveness suggests that the intervention was most effective for those with a lower issue familiarity, but this was also not significant (β = 0.15, p = .25).

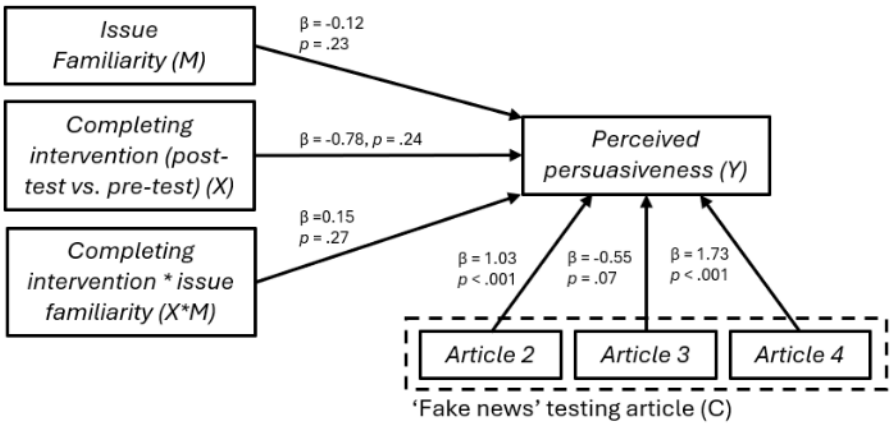


Figure 8: Moderation analysis of the effect of the intervention on perceived 'fake news' persuasiveness, and how this relationship is moderated by issue familiarity.

Although inoculation theory has largely been considered a cognitive and affective intrapersonal process, the affective elements have historically received less attention from communication scholars. Scholars outside of inoculation theory have argued that emotion is a fundamental part of the process of persuasion (Dillard, 1998), and studies investigating boundary conditions of inoculation theory have time and again returned to affect (Pfau et al., 2001). Initial research with the intervention *Fake News* saw affect as an important variable and measured individuals' affective responses to the intervention through sentiment analysis of open-ended questions, finding that the treatment group displayed greater affective content. A similar analysis was completed in this study, comparing the affective content between the three testing stages and the online control. Python's Natural Language Toolkit VADER sentiment analysis model was used in this analysis (Hutto & Gilbert, 2014). A score was awarded for individuals' sentiment in all three open-ended questions (see supplementary materials for full breakdown of this process). A score was awarded to each individual survey response, representing the collective negative sentiment from all three survey questions.

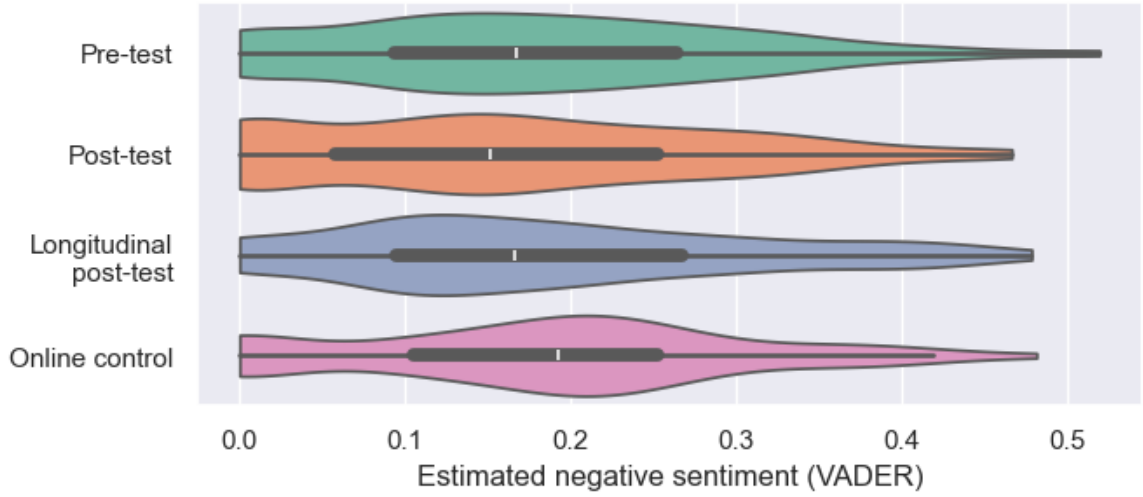


Figure 9: Estimated negative sentiment of responses to open ended questions on 'fake news' articles, between three within-subjects stages and an online control group.

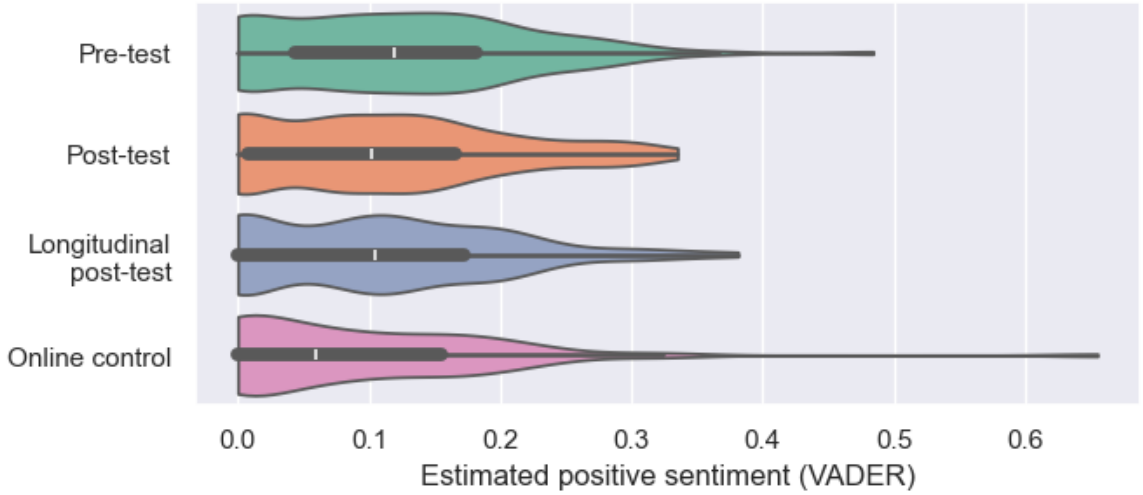


Figure 10: Estimated positive sentiment of responses to open ended questions on 'fake news' articles, between three within-subjects stages and an online control group.

Affective content (both positive and negative sentiment scores) was descriptively lower after completing the intervention, trending against the expected direction. There were no significant changes in sentiment between the pre-test, post-test, and longitudinal post-test ($H(2) = 1.35$, $p = .51$; $H(2) = 1.88$, $p = .39$ for positive and negative sentiment respectively). It is notable that the online control group displayed descriptively more negative and less positive sentiment in open-ended questions than those completing the intervention. For the online control group, there was a significant correlation between the negative sentiment displayed through open-ended survey questions and the perceived persuasiveness, agreement, and reliability of 'fake news' article ($r(73) = -0.25$, $p = .031$; $r(73) = -0.23$, $p = .044$; $r(73) = -0.33$, $p = .004$ for persuasiveness, agreement, and reliability measures respectively). That is, individuals in the online control group who expressed negative sentiment in open-ended survey questions also perceived the 'fake news' article to be less persuasive, reliable, and agreed with it less. This was a relationship not observed in the groups who completed the

intervention. This suggests that, perhaps, individuals completing the survey in-person may have felt similar emotional responses which guided their quantitative judgements, but did not feel able to do so while supervised by the researcher.

This section has summarised the quantitative results from three measures in this study: perceived familiarity, perceived agreement, and perceived reliability of 'fake news' articles from three testing stages in this study, as well as an online control group. The intervention was found to have no significant effect on individuals' resilience to misinformation, nor between the longitudinal post-test and the online control group; however, perceived persuasiveness judgements did descriptively reduce after completing the intervention. Two moderating variables central to the inoculation theory are also tested: issue familiarity and affect. Although not significant, the effectiveness of the intervention was descriptively reduced the higher individuals' issue familiarity. The affective content evaluated from sentiment analysis of open-ended questions did not significantly change between testing stages, although it is notable that the affective content in the online control group was significantly more negative. Important to highlight is the impact floor effects may have had on the three main measures (i.e., perceived persuasiveness, reliability, and agreement), which may be partly responsible for the null results observed.

The intervention was descriptively most effective for those with a low issue familiarity, differing to the curvilinear relationship suggested in inoculation theory literature. This points to the influence of floor effects on issue familiarity; as immigration is a central and polarising subject in British politics (Blinder & Allen, 2016), and it is unlikely that an individual could be entirely unfamiliar with this issue. Thus, one may conclude that participants in this study are moderately or highly familiar with the topic. The significant difference of affective content from open-ended measures between the longitudinal post-test and the online control suggest that these open-ended questions are susceptible to mode effects. More specifically, individuals having taken part in the in-person study answered open-ended questions with more conservative language than the online control group, even when the survey was completed online. This mode-effect has potential implications for further research using open-ended survey questions on item evaluation tasks between survey modes. In the following section I compare some of the findings presented in this section with previous research with the same intervention by Roozenbeek and van der Linden (2018).

4.2 Adults Versus Adolescents: A Replication of Roozenbeek and van der Linden (2018)

As well as a within-subjects longitudinal mixed-methods research in its own right, this study is also a replication of previous research with the same intervention by Roozenbeek and van der Linden (2018). Replication research is essential as it can help give ecological validity and generalisability to previous findings and can provide nuance in context or audience for which the effects of the original research may differ. Individual scientists, however, are incentivised to prioritise novelty over reproducibility (Aarts et al., 2015), leading to claims of a 'replication crisis' (Nosek et al., 2025). Within game-based inoculation research scholars have at times struggled to replicate previous findings (Roozenbeek et al., 2024). For example, research by M. E. Graham et al. (2023) conducted a replication study with the game-based intervention *Bad News*, findings that the intervention reduced belief in not only false Tweets, but true Tweets to (often) the same extent. Furthermore, Modirrousta-Galian and Higham (2023) conducted reanalyses of game-based misinformation interventions *Bad News* and *Go Viral!*, finding that the interventions did not improve discrimination between true and false information.

The present study and research by Roozenbeek and van der Linden (2018) share many similarities that make this suitable as a replication. Firstly, both studies use the same intervention, *Fake News*. Secondly, this study collected the same socio-demographic variables as the original study (i.e., political ideology, age, issue familiarity, and gender), as well as additional information on nationality and education. This study also included very similar 'fake news' articles as the original study[11] with the same quantitative (perceived persuasiveness, personal agreement, and perceived reliability) and qualitative (open-ended questions asking participants to highlight arguments they find persuasive, reliable, and what they think the writer is trying to convince them of) measures. This study has three key differences to the original by Roozenbeek and van der Linden (2018). Firstly, as this study included a longitudinal element two further articles were created, although analysis in this section does not include the longitudinal post-test, and results only include responses from the first two 'fake news' articles. Secondly, this study is within-subjects (rather than the original, which was between-subjects). The difference between the original treatment group and the immediate post-test (T2) in this study is that individuals completing T2 will have also completed the immediate pre-test (T1), a difference that is likely to increase the effectiveness of the intervention (see testing effects, Maertens et al., 2021; Roozenbeek

[11] The two articles used from previous research by Roozenbeek & van der Linden (2018) featured very minor adaptations in this study such that they suited a UK audience.

et al., 2024). Finally, original research by Roozenbeek and van der Linden (2018) involved adolescents at a Dutch public high school, a far younger participant pool than this study which involved individuals from 18 to over 60 years old. Thus, this study is a replication with extension (Brandt et al., 2014).

Comparing results from the three core quantitative measures (i.e., perceived persuasiveness, reliability and personal agreement) to other research with *Fake News* with adolescents shows a higher prior knowledge and awareness of misinformation in adult participants (versus adolescents). Overlaying results from previous research (Roozenbeek & van der Linden, 2018) with this intervention shows that the young adolescents in the original study had a higher baseline susceptibility to misinformation, finding 'fake news' articles more persuasive, reliable, and agreeing with them more than the adult participants in the present study (see Figure 11). A difference between the two studies is that the original study had a largely younger participant group (public high school students), which may be responsible for this difference.

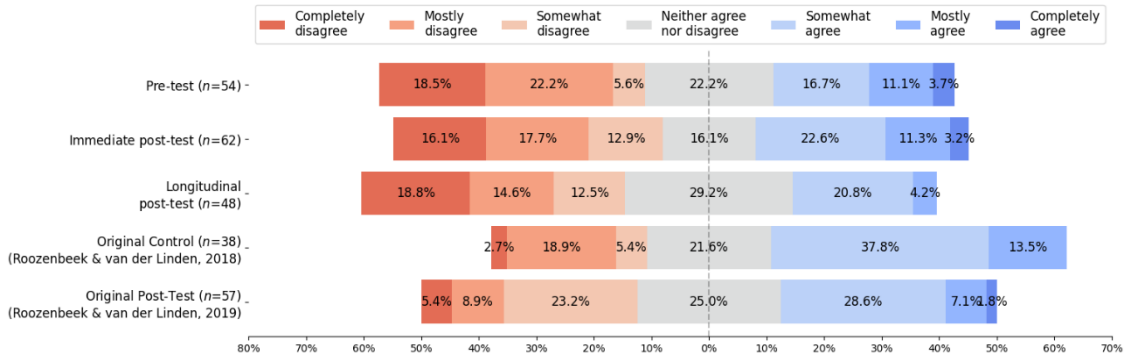


Figure 11: Participant agreement judgements of 'fake news' articles 1 and 2 across all three testing stages. Also shown are control and treatment groups from original research by Roozenbeek and van der Linden (2018). Only judgements on articles 1 and 2 from the present study are included as these were used in the original research by Roozenbeek and van der Linden (2018) and are 'most-recent' undersampled. Data is given as percentages to scale the longitudinal post-test (after participant attrition) and data from original research by Roozenbeek and van der Linden (2018).

Findings from Roozenbeek and van der Linden (2018) found a significant difference in perceived reliability of the 'fake news' article. The treatment group rated the 'fake news' article to be significantly less reliable than the control group (M = 3.06, SE = 0.20 vs. M = 3.60, SE = 0.27, $t(90) = 1.69$, $p < .05$, one-tailed, Cohen's d = 0.36). When considering results from only the 'fake news' articles used in the original experiment, this study instead finds an increase in perceived reliability of the 'fake news' article (M = 2.48, SE = 0.21 vs. M = 3.03, SE = 0.19, $t(106) = -1.92$, p = .06, two-tailed, Cohen's d = 0.36), although this result is not significant. Similarly to original research (M = 3.67, SE = 0.21 vs. M = 3.97, SE = 0.21, $t(93) = 1.00$, p = .16, one-tailed, Cohen's d = 0.21), this study also finds that perceived persuasiveness of 'fake news' articles are descriptively lower after taking part in the intervention (M = 3.52, SE = 0.27 vs. M = 3.44, SE = 0.21, $t(106)$ = 0.25, p = .40, one-tailed, Cohen's d = 0.05). Finally, where original research with

adolescents finds that the intervention descriptively reduced personal agreement with 'fake news' articles (M = 3.91, SE = 0.18 vs. M = 4.14, SE = 0.23, t(91) = 0.77, p = 0.22, Cohen's d = 0.16), the present study found the intervention descriptively increased personal agreement with 'fake news' articles (M = 3.44, SE = 0.25 vs. M = 3.58, SE = 0.23, t(106) = -0.41, p = .68, two-tailed, Cohen's d = 0.08).

Further analysis in original research with adolescents involved a mediation analysis, finding that reliability mediated the effect of the intervention on perceived persuasiveness measures (see Figure 12). That is, the higher the judged reliability, the more persuasive people found the article. Following the same procedure as Roozenbeek and van der Linden (2018), we completed a mediation analysis to investigate the mediating effects of perceived reliability on the relationship between completing the intervention and perceived persuasiveness of 'fake news' articles. As in the original analysis, age, gender, political ideology, and issue familiarity were included as covariates. Only judgements on articles one and two were included, the effect was only measured between the pre-test (T1) and immediate post-test (T2), and standard errors were bias-corrected and bootstrapped 1,000 times. The mediation analysis from the present study (see Figure 13) found that although the higher perceived reliability of 'fake news' the higher the perceived persuasiveness (β = 0.76, p < .001), similarly to original research. However, completing the intervention did not significantly impact perceived 'fake news' article reliability, diverging from original research by Roozenbeek and van der Linden (2018). That said, playing the intervention did have a significant indirect on perceived persuasiveness (β = -0.56, p = .03), more so that original research with adolescents (β = -0.38, p = .034).

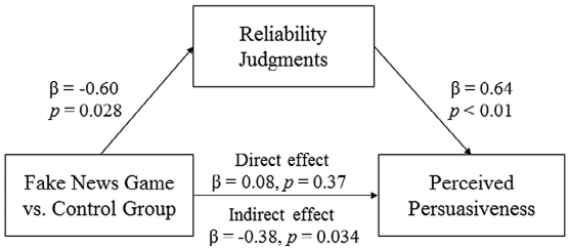


Figure 12: Mediation model from original research with adolescents by Roozenbeek and van der Linden (2018). Coefficients are unstandardised, and covariates include gender, age, issue familiarity, and political ideology.

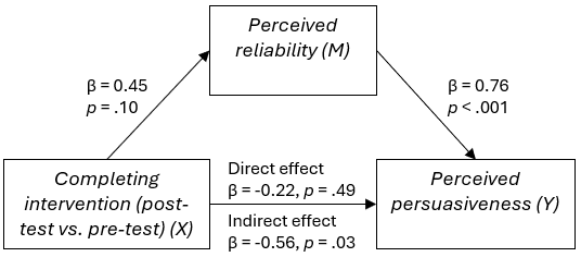


Figure 13: Mediation model from the present study, using judgements from 'fake news' articles 1 and 2 only. Coefficients are unstandardised, and covariates include gender, age, issue familiarity, and political ideology.

A further mediation analysis can be completed that includes ratings on not only the two 'fake news' articles used in original research by Roozenbeek and van der Linden (2018), but all four articles used in the present study (see Figure 14). With judgements from all four articles included, the analysis is afforded a higher power. Although perceived reliability still predicts perceived persuasiveness (β = 0.79, p < .001),

the indirect effect of the intervention on perceived persuasiveness judgements is no longer significant (β = -0.23, p = .20). It is possible that the two extra 'fake news' articles written for the present study (used in conjunction with the original articles by Roozenbeek & van der Linden, 2018) differed in such a way that perceived reliability, persuasiveness, and participant agreement was harder to shift.

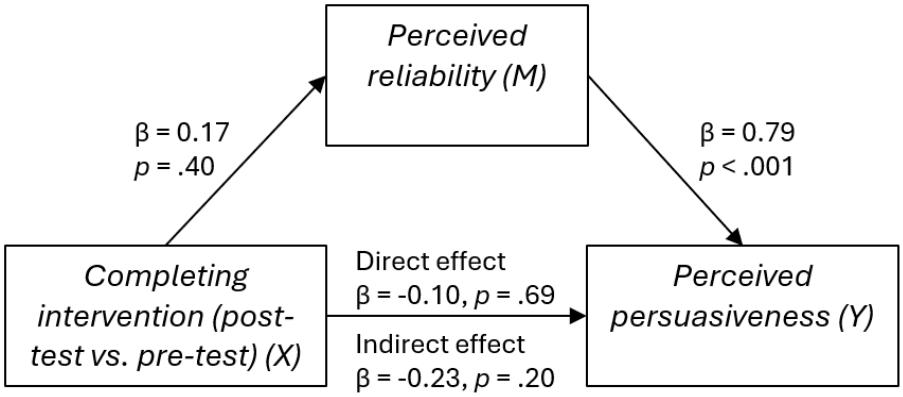


Figure 14: Mediation model from the present study, using judgements from all 'fake news' articles. Coefficients are unstandardised, and covariates include gender, age, issue familiarity, and political ideology.

This section has presented the present study through a different lens: as a replication of previous research with the same intervention (Roozenbeek & van der Linden, 2018). Specifically, this section has summarised analyses comparing the effectiveness of the intervention and mediating properties of perceived reliability on perceived 'fake news' article persuasiveness.

A good replication study attempts to emulate as similar conditions believed to be necessary to replicate previous results (Aarts et al., 2015; Schmidt, 2009). A key difference between the original research by Roozenbeek and van der Linden (2018) and the present study is socio-demographic differences between participants. All participants are at a very similar level of age and education, whereas participants in the present study varied drastically on both of these measures. Although speculative, one can assume that the high-school adolescents were also already familiar with one another before completing the intervention, whereas almost all participants in the present study were meeting one another for the first time. Nonetheless, this was by design. The present study was motivated to not only replicate previous findings, but to further explore how effects maintained effectiveness over time, and to explore these effects with a wider audience.

The present study did not replicate the previous findings (Roozenbeek & van der Linden, 2018). Where previous findings show that the intervention significantly decreased perceived reliability of 'fake news' articles, the present study find the opposite to be true with the same articles (although not significant). Similarly, while previous findings demonstrate how perceived reliability can mediate the effect of the

intervention on perceived persuasiveness, a mediation model from the present study (including all four articles, see Figure 14) found this indirect effect to not be significant. Nonetheless, the replication did give insights into how the intervention impacted individuals of different awareness and experience of misinformation. The control group in the original study perceived 'fake news' articles to be descriptively more persuasive ($t(108) = -1.57$, $p = .12$), and significantly more agreeable ($t(108) = -2.23$, $p = 0.03$) and reliable ($t(108) = -2.62$, $p = .01$) than the immediate pre-test (T1) in the present study. This suggests that the adolescents in the original study exhibited a higher baseline susceptibility than those in the present study. This is expected, as adolescents in the original study may still have been learning about the modern information landscape. It is for this reason, perhaps, that the original study showed a more significant reduction in scores in the post-test. Where the intervention was likely prophylactic for the young adolescent participants in this original study, the intervention was likely more therapeutic for the adult participants in the present study.

Despite this, there were some notable caveats to this replication. The present study found that the articles themselves acted as significant and impactful covariates, particularly if the political leaning of the article (i.e., liberal or conservative) aligned with the ideological leaning of the individual. Unfortunately, these variables were not recorded in the previous study (Roozenbeek & van der Linden, 2018), and as such this interaction cannot be explored in this replication. Furthermore, session-level variables that have given insight into the mechanisms and dynamics of each intervention (see previous Section) were not included either.

4.3 Demographic Variables Moderating Intervention Effectiveness

Despite continued interest and an encouraging meta-analysis (Banas & Rains, 2010), inoculation has not found universal success. Research on human behaviour and psychology can often make ecological assumptions from findings that focus on Western, educated, industrialised, rich and democratic (WEIRD) populations, that simply do not transfer to non-WEIRD societies (Henrich et al., 2010). Misinformation inoculation research has seen comparatively little interest in non-WEIRD populations (Roozenbeek et al., 2024), and game-based inoculation studies have in some cases found no significant impact with these populations (Harjani et al., 2023; cf. Cook et al., 2024). Although not the focus of this study, socio-demographic variables such as nationality were collected and have allowed investigation into their moderating effect. This can help build a picture as to why an intervention may have (not) been effective and identify intervention-resistant demographics where further thought is required on intervention design.

Following from the previous section, I present and discuss some promising socio-demographic findings from the study. In particular, this section assesses how individuals' nationality, age, gender, education level, and political ideology impact perceived persuasiveness, reliability, and personal agreement with 'fake news' articles, and how these socio-demographic variables moderated the effectiveness of the intervention *Fake News*. This section also offers a discussion of implications from socio-demographic findings. Further statistical analyses with socio-demographic variables are provided in the appendix (Appendix 1; Appendix 4).

Political ideology and polarisation are important variables in modern misinformation research. Political polarisation has been shown to be a key variable in misinformation on topics such as climate change (Lewandowsky, 2025) and vaccine hesitancy (J. H. Williams et al., 2023). Analyses of political ideology's direct effect on measures can give insight into how individuals with conservative or liberal ideologies perceive the persuasiveness and reliability of 'fake news' articles. Individuals stated their political ideology on a scale from 0 (labelled 'Labour') to 100 (labelled 'Conservative') at the beginning of the study, and each testing phase response (i.e., T1, T2, and T3) was labelled as whether it aligned with their ideological leaning (depending on whether the 'fake news' article was left leaning or right leaning. As expected, individuals perceived 'fake news' articles to be more persuasive ($t(328) = 7.20$, $p < .001$, Cohen's $d = 0.87$), reliable ($t(328) = 6.33$, $p < .001$, Cohen's $d = 0.77$), and agreed with them more ($t(328) = 11.30$, $p < .001$, Cohen's $d = 1.37$) when they aligned with their own political ideology.

A hierarchical multiple regression analysis was also conducted to test if the article's ideological alignment to the reader moderated the impact of the intervention on perceived reliability ($R^2 = .230$, $F(6, 191) = 9.49$, $p < .001$). To negate the impact of an unbalance of testing articles post-test vs. pre-test, these were added as covariates (see Figure 15). The intervention was descriptively more effective at reducing perceived reliability of 'fake news' articles when these articles were ideologically unaligned, although this effect was not significant ($\beta = -0.66$, $p = .10$). This suggests that completing the intervention improved (reduced) perceived reliability of 'fake news' articles that did not align with individuals' personal ideology; however, 'fake news' articles that did align ideologically were more resistant to the inoculation intervention.

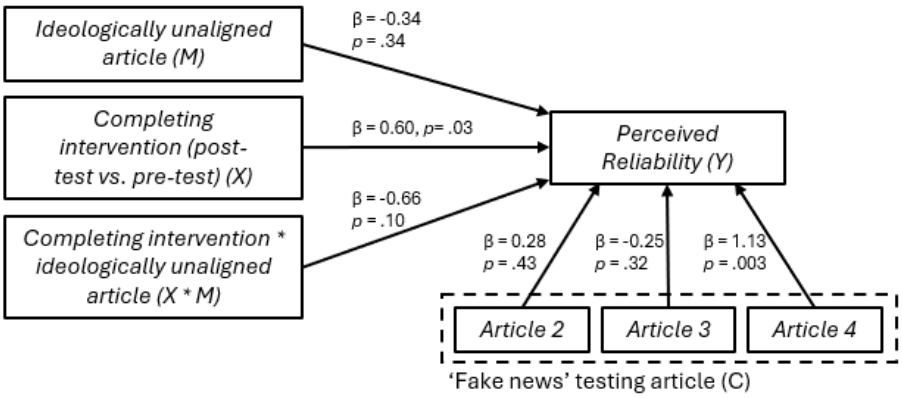


Figure 15: Statistical diagram of a moderation analysis, testing the effect of the intervention Fake News on perceived reliability of 'fake news' articles, and the ideological alignment in relation to the individual moderated this relationship. Categorical covariate is individual articles.

As only 19 participants (16.1%) declared their nationality as being anything other than British, these were coded together as "non-British nationality"[12]. Although we are not able to investigate variations in non-WEIRD nationalities individually (due to a lack of power), this does give insight in how an intervention with a topic on immigration can confer resistance for individuals who are not based in their home nationality. Individuals with a non-British nationality perceived 'fake news' articles to be significantly more persuasive ($t(328) = -3.60$, $p < .001$, Cohen's $d = -0.54$), reliable ($t(328) = -3.20$, $p = .001$, Cohen's $d = -0.48$), and agreed with them more ($t(328) = -2.59$, $p = .01$, Cohen's $d = -0.29$). Nationality also moderated how the *Fake News* intervention affected individuals' perceived persuasiveness of 'fake news' articles. A hierarchical multiple regression analysis was conducted to test how nationality moderated the impact of the intervention on perceived persuasiveness ($R^2 = .26$, $F(6, 229) = 13.55$, $p < .001$). To negate the impact of an unbalance of testing articles post-test vs. pre-test, these were added as covariates. Examination of the interaction between the intervention and nationality revealed that the intervention had been descriptively more effective in reducing perceived persuasiveness judgements for individuals with non-British nationalities ($\beta = -0.88$, $p = .14$). Despite this relationship not being significant, post-hoc tests were completed to better understand this relationship; specifically, if nationality's impact on measures was mediated by another variable (e.g., issue familiarity, education, or political leaning). This post-hoc analysis did not reveal significant mediating variables.

[12] We recognise that this coding approach and statistical analyses lacks a certain degree of precision and power, and stress to the reader to treat these results as 'exploratory'.

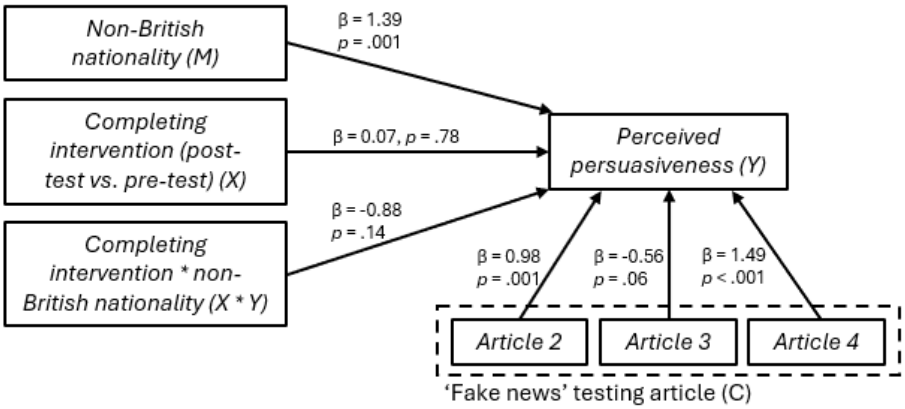


Figure 16: Statistical diagram of a moderation analysis, testing the effect of the intervention Fake News on perceived persuasiveness of 'fake news' articles, and how nationality moderates this relationship. Categorical covariate is individual articles.

Additional Insights

Previous studies have suggested that demographic variables like age, gender, and education level affect views on misinformation and immigration (Blinder & Allen, 2016; Kyrychenko et al., 2025), so further tests were completed on these variables. These analyses are important as, if individuals with a lower age or education level perceived 'fake news' articles to be more persuasive, reliable, or agreed with them more, this may be a demographic that warrants particular attention for future curated interventions.

Gender had no direct or moderating effect on outcome measures, but age and political ideology were found to have significant direct effect. Older individuals perceived 'fake news' articles as being less persuasive ($r(326) = -0.19$, $p < .001$), reliable ($r(326) = -0.21$, $p < .001$), and agreeable ($r(326) = -0.18$, $p = .001$). This aligned with recent research investigating socio-demographic factors that impact misinformation susceptibility (Kyrychenko et al., 2025). Individuals with conservative political ideology also had a higher base susceptibility to misinformation, further aligning with research by Kyrychenko et al. (2025). Political ideology was significantly correlated to perceived persuasiveness ($r(327) = 0.19$, $p < .001$) and reliability ($r(327) = 0.13$, $p = .02$) and descriptively correlated to agreement ($r(327) = 0.10$, $p = .08$) with 'fake news' articles. However, neither age nor political ideology did not significantly moderate the effectiveness of the intervention. Although education level did not have a significant direct effect on perceived persuasiveness ($r(326) = -0.04$, $p = .48$), reliability ($r(326) = -0.07$, $p = .22$), or agreement ($r(326) = 0.005$, $p = .93$) with 'fake news' articles, it did significantly moderate the effect of ideologically alignment of 'fake news' articles across the three measures. Put simply, individuals with a higher education had higher judgements across all three measures for 'fake news' articles that were ideologically aligned with them, and had lower judgements across measures for 'fake news' articles that did not ideologically align, than those who were less educated (see Figure 17).

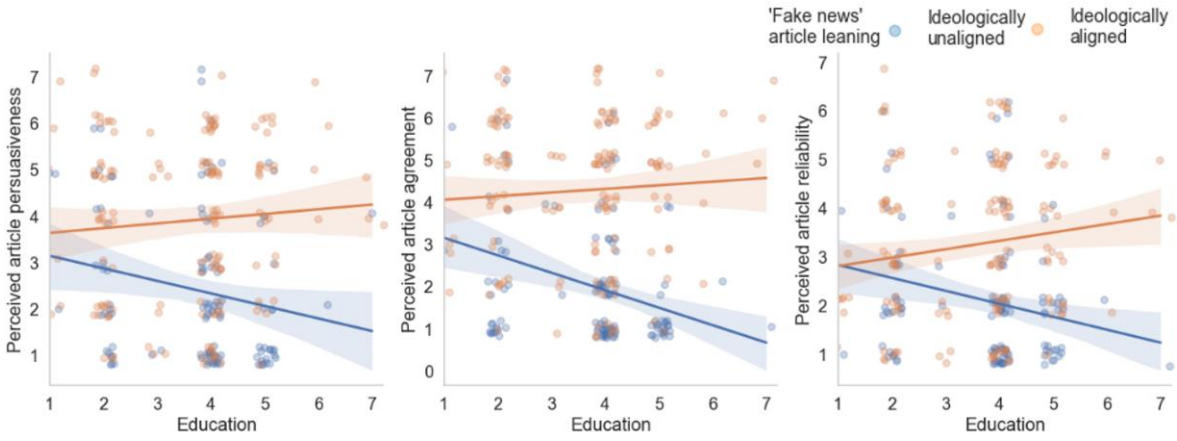


Figure 17: Comparison of the effect of education level on persuasiveness, agreement, and reliability measures between ideologically (un)aligned 'fake news' articles.

This section has summarised the direct and moderating effects of socio-demographic variables on outcome measures. Specifically, the impact of political alignment, education, age, gender, and nationality has been assessed on perceived persuasiveness, reliability, and agreement with 'fake news' articles. How this relationship changes after completing the intervention is also assessed.

Participant responses were coded depending on if the 'fake news' articles were aligned to their own political ideology, and as expected individuals found ideologically aligned 'fake news' articles more persuasive, reliable, and agreed with them more (the reverse was also true of unaligned 'fake news' articles). Although not significant, ideological alignment of the article descriptively moderated the intervention's effect on perceived reliability. That is, the intervention was descriptively more effective at reducing perceived reliability of ideologically unaligned 'fake news' articles (vs. aligned). The higher resilience of ideologically aligned 'fake news' articles is expected. Scholars have long studied how individuals are not impartial readers of new information but accept or reject information depending on their prior beliefs on the topic (Bolsen et al., 2014; Kunda, 1990; Osmundsen et al., 2021; Van Bavel et al., 2021). Given this, one can explore then results from the first section of this chapter under a new lens. It is not that the game-based intervention *Fake News* was entirely ineffective at building resilience to misinformation, but that perceived persuasiveness of ideologically aligning 'fake news' articles are especially difficult to shift.

Although not significant, it was also found that nationality descriptively moderated the effect of completing the intervention. That is, individuals with non-British nationality perceived 'fake news' articles to be less persuasive after completing the intervention. This result could be interpreted in two ways, a) the design of this intervention (i.e., collaborative, in-person) could be particularly well designed to individuals with a non-British nationality, or b) individuals with a British nationality are more resistant to interventions aiming to improve their resilience to

misinformation. Due to sample sizes, the coding of this demographic variable was binary (i.e., as either having British nationality, or not). Regrettably, the collapsing of non-British nationalities into a single value negated much nuance and variation between nationalities, and did not allow a comprehensive analysis of non-WEIRD populations (Henrich et al., 2010)—analysis that is sorely needed in the field (Roozenbeek et al., 2024).

This section also analysed the direct effect of socio-demographic variables in outcome measures. Recent research by Kyrychenko et al. (2025) has sought to paint a clearer map of socio-demographic characteristics that are more common among those vulnerable to misinformation. In this wide-ranging study, Kyrychenko et al. (2025) measure susceptibility using the Misinformation Susceptibility Test (MIST), where participants rate the veracity of headlines. They conclude that misinformation susceptibility is higher in younger, non-male, less educated individuals. Findings from this study align with this wider research by Kyrychenko et al. (2025), finding that non-British, ideologically conservative, younger individuals find long-form textual 'fake news' articles to be more persuasive, reliable, and agreeable. In the following section, I consider how the design of the game-based intervention *Fake News* moderated its effectiveness.

4.4 Role of Intervention Experience

Research has demonstrated that game-based inoculation theory interventions can be an effective individual-level method to building resistance to misinformation in players (Kiili et al., 2024). However, game-based interventions are often seen as 'black-box' activities (i.e., closed systems that can only be viewed in terms of inputs and outputs), with little known about how interventions increase resistance (beyond inoculation mechanisms). As such, scholars have at times been unsure why some interventions have not been effective (Appel et al., 2025; Harjani et al., 2023). By increasing transparency of how players interact with the intervention, practitioners may be better informed of how game mechanics affect the intervention experience, and how individuals from different socio-demographic backgrounds interact with design elements differently. This can lead to not only more effective interventions but may help practitioners create interventions specific to target groups.

This study has sought to give more insight into the mechanisms taking place in the intervention *Fake News* by recording variables such as group size, interaction level, group score, and the character the group played as. In this section, I present and discuss these session-level variables, the insight they give into the dynamics of groups and groups' interaction with the intervention, and if they affected how the intervention conferred resistance in players. The section starts with an assessment of the moderating

effects of these session-level variables is assessed, and how the character that each group played as changed the experience of the intervention. Additional insights on group size and interaction level are also discussed.

The level of group interaction, the score achieved, the character played as, and the group size had no significant direct or moderating effects on intervention outcomes. Four hierarchical multiple regression analyses were completed to explore the moderating effect of group interaction (R^2 = 0.256, F(6, 121) = 6.944, p < .001), group score (R^2 = 0.221, F(6, 227) = 10.76, p <.001), character played as (R^2 = 0.290, F(10, 117) = 4.780, p < .001), and group size (R^2 = 0.224, F(6, 227) = 10.89, p < .001) on perceived persuasiveness of 'fake news' articles. In each of these analyses, the overall ratings on individual 'fake news' articles were added as covariates, which were mostly responsible for high R^2 values. The intervention was descriptively more effective for groups who interacted more during the session (β = -0.29, p = .49), but this result was insignificant (see Figure 18). Group score had little moderating impact on the effect of the intervention (β = 0.04, p = .61), and this result was also not significant. The moderating impact of the group size on intervention effectiveness at reducing perceived persuasiveness of 'fake news' articles was also insignificant (β = -0.22, p = .52).

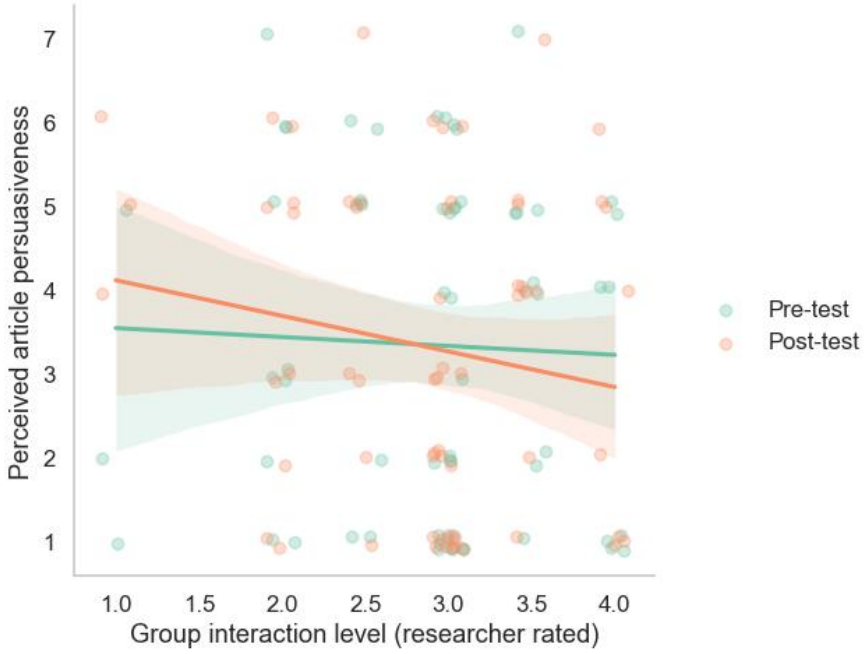


Figure 18: Perceived persuasiveness of 'fake news' articles pre- and post-intervention, and how this is moderated by the level of inter-group interaction during the intervention.

The lack of moderating power from in-session variables suggests that the intervention may have been ineffective for all groups, irrespective of group size, level of interaction during play, group score, or the character the group played as. To further explore this, further analysis can be completed on how the intervention design differed between groups. At the beginning of each session, groups were assigned one of four characters: Dennis in Denial, Mike the Money Changer, Paula for Political Posturing, or

Dana and Fox: Co-Conspirators. Each character had their own background, motivation, and method which groups had to consider when writing their character's article. For example, Paula for Political Posturing was motivated by political gain and used hyperbole, ad hominem attacks, and common-man appeal as methods for influencing public opinion. Therefore, each group focussed on and practiced use with different 'fake news' tactics (Roozenbeek & van der Linden, 2018). Although no session-level variables significantly moderated the effect of the intervention, they were correlated with one another. The following paragraphs further explore these variables, drawing on game design literature.

Research with collaborative analogue games suggests that game difficulty is an important variable in increasing engagement (Zagal et al., 2006). Matching game difficulty to players' expertise ensures that players work collaboratively to complete tasks by sharing knowledge and resources. Group score differed between characters, suggesting that some game characters' misinformation techniques were harder to identify when completing the intervention task. Figure 19 illustrates that groups playing Dennis in Denial, for example, scored far higher on average than other characters. A one-way ANOVA was conducted to measure score differences between characters and found this difference to be significant ($F(4, 64) = 3.77$, $p = 0.03$). Further analysis found that group interaction did not significantly change depending on the character the group played as ($F(4, 64) = 0.13$, $p = 0.94$). That is, groups did not interact more or less depending on which character they had been assigned, and by extension groups did not interact significantly more or less if the tasks were harder or easier.

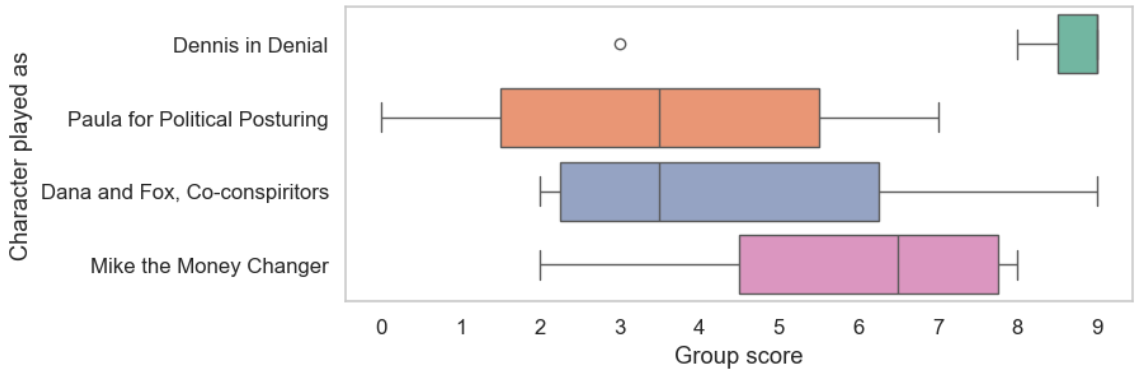


Figure 19: Scores for groups playing as different characters in the game-based intervention Fake News. Results do not include round 1 of data collection.

Further analysis was completed on group interaction, and if this was directly affected by other session-level or socio-demographic variables. Group size was found to have no significant impact on interaction level, with larger groups having only a descriptively higher level of inter-group interaction during the intervention ($r(24) = .13$, $p = .54$, see Figure 20). Moreover, group score had no significant effect on score achieved during the session ($r(24) = -.06$, $p = .76$). With regards to socio-demographic impacts on

session-level variables, interaction level was most impacted by age, although this was not significant. Individuals with a higher age had descriptively higher-interacting groups ($r(63) = .16$, $p = .19$), but this correlation was not significant. As only the second round of data collection included the recording of character played and interaction level, many of these tests lacked power.

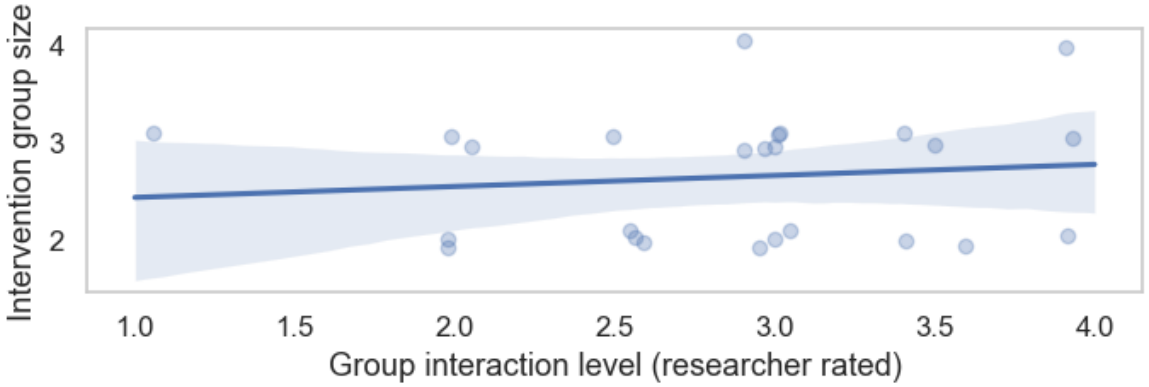


Figure 20: Effect of group size on level of interaction observed during a session, and the effect of group size on recorded group score.

This section has summarised the moderating effects of session-level variables on intervention effectiveness and sought to explore the internal mechanisms of the intervention. Specifically, this section has analysed the impact between group size, interaction level, group score, and character group played as, and interaction with socio-demographic variables. Although none of the session-level variables had a significant moderating impact on the intervention, many trended in the expected direction. The intervention was descriptively more effective at reducing perceived persuasiveness in 'fake news' articles when the group had interacted more when completing the intervention, and the intervention was descriptively more effective for groups with lower scores.

Treating the intervention as a 'white box' system has also allowed insight into game difficulty, interaction, and experience. Average scores for different characters differed significantly suggesting differences in difficulty between them. However, players are not recorded interacting significantly more when their group played as a 'difficult' character. This is unexpected, as it is intuitive that players must interact more when working collaboratively to solve a harder problem. This is reliant on a feedback loop to inform the player of when they are completing a task (un)successfully. For example, the board game Pandemic requires players to work collaboratively as different disease-fighting specialists to prevent the world from succumbing to four viral diseases. As more of the world is overcome with disease the game becomes increasingly more difficult, and players must work increasingly collaboratively to devise strategies to overcome this (Anania et al., 2016). In *Fake News*, players only become aware of their score after completing the intervention, with very little feedback of success during play.

This may go some way to explain why interaction did not differ between groups playing characters of different difficulty.

It was also unexpected that higher group sizes did not result in more inter-group interaction during the intervention. I propose that this could be due to two reasons: social loafing and intergroup anxiety. Participants were financially compensated for taking part in the study and were not penalised for low engagement or interaction with other group members. Social loafing suggests that individuals have a tendency to expend less effort when working on a task collaboratively (versus individually) (D. W. Johnson & Johnson, 2009; Karau & Williams, 1993; Pandeirot & Aseng, 2017). Perhaps, then, as the intervention does not require cooperation from all players (e.g., if players had their own private cards, the intervention may not be completable without their cooperation, see Zagal et al., 2006), some individuals in larger groups exhibit lower motivation to interact with other group members. Insignificant correlation between group size and interaction can also be explained by intergroup anxiety. That is to say, individuals may feel anxious to raise opinions on in-game choices when working with people from different socio-demographic groups (Stephan, 2014; Stephan & Stephan, 1985). As players were members of the public and did not know one another before the intervention, and that the intervention itself focussed on a polarising topic in British politics (i.e., immigration) it is likely that intergroup anxiety played a role in larger intervention groups. That said, further research is required to establish causality of these variables on the intervention.

Summary and Conclusion

Game-based misinformation interventions are a promising individual-level approach to build resistance to misinformation in individuals, and research has demonstrated that this approach can be effective in reducing vulnerabilities to misinformation on topics such as climate change (Cook, 2021), health campaigns (Appel et al., 2025), and politics (Roozenbeek & van der Linden, 2020). Despite these promising results, not all interventions have demonstrated success (Harjani et al., 2023), and there is a lack of evidence on how effects decay over time (Roozenbeek et al., 2024), especially on different platforms (Kiili et al., 2024). This chapter has sought to present quantitative results from research completed in 2024 with the analogue game-based inoculation intervention *Fake News*. This chapter has aimed to answer the first research question of this thesis: how effective are analogue game-based interventions against misinformation, and to what extent does this change over time.

This chapter has presents results from four different perspectives: as an inoculation theory intervention, as a replication study, as an analysis of the role of socio-demographic variables, and as a 'white box' system. As an inoculation theory

intervention, I have not only presented results from the study's base measures but also presented and discussed the moderating effect of issue familiarity and affect, important variables in the inoculation theory literature (Banas & Rains, 2010; Compton, 2012; Pfau et al., 2005). This has allowed me to understand if the intervention has not only been effective, but if it has stimulated key inoculation mechanisms as expected. Next, results from the present study are considered as a replication of Roozenbeek and van der Linden (2018), reflecting on differences of sample groups between the two studies. Following this, the chapter has offered results pertaining to individuals' socio-demographic traits, giving insight into not only if the sample was population-representative but whether the intervention was effective some nationalities, ages, and political ideologies than others. Finally, findings from session-level variables give insight into how the intervention builds resistance in players. The intervention is not treated as a closed, 'black box' intervention (i.e., a system that is considered in terms of inputs and outputs), but as an open activity whose game-play mechanisms can be understood. In answering research question 1, this chapter offers two contributions to the game-based inoculation literature.

Firstly, the quantitative results from this study represent the second unsuccessful replication of original research by Roozenbeek and van der Linden (2018) using the same intervention and measures (Wong & Wu, 2023) and call for a greater scrutiny of intervention design and evaluation methods. There has been considerable interest in replicating interventions such as *Bad News* and *Go Viral!* specifically in new contexts and with new populations (Axelsson et al., 2024; Basol et al., 2020; Roozenbeek et al., 2020; Traberg et al., 2024; Wong & Wu, 2023). However, replication research has mostly focussed on the intervention *Bad News*, ignoring harder-to-study in-person interventions such as that used in the present study. These findings come at a time when recent replications of other game-based interventions have failed (M. E. Graham et al., 2023; Wong & Wu, 2023) or may have even had a negative overall impact (Modirrousta-Galian & Higham, 2023). I have argued that these null results may be caused by floor effects in quantitative measures, the low level of threat and high level of issue familiarity assessed in participants. I also note that the quantitative measures used in this study are narrow in scope (focussing on changes in player behaviour), and do not include player transformations of knowledge, skill, affect, or belief (see 'types of transformations' Culyba, 2018, p. 91). That this is the second failed replication of Roozenbeek and van der Linden (2018) places scrutiny on this initial research, and places further scrutiny on this original research.

The second contribution that these quantitative results make to the literature is to highlight the importance of prior issue familiarity in intervention design. Although not significant, the effect of the intervention on perceived persuasiveness,

reliability, and personal agreement with 'fake news' articles trended in the expected direction (i.e., reduced after completing the intervention, when controlling for covariates). Importantly, individuals' familiarity with the intervention's topic moderated this relationship such that the intervention was more effective for those with a lower issue familiarity. Inoculation theory highlights prior knowledge as an important variable in the inoculating process: too high, and the individual is already aware of their attitudinal vulnerability (Compton, 2012; Pfau et al., 1997, 2005). Findings in this chapter demonstrate that adult participants in this study demonstrated a significantly higher baseline resilience to 'fake news' articles than adolescents in previous research (Roozenbeek & van der Linden, 2018). I hypothesise, therefore, that the high issue familiarity rendered the inoculation intervention ineffective for the study's sample. This further highlights the importance of prior familiarity as a variable in inoculation literature, and one that demands consideration by practitioners designing new interventions.

Despite the important findings and promising insights presented in this chapter, results raise new open questions in the field of game-based inoculation, particularly around the participants themselves, and how they interact with the intervention. Given that original research with adolescents (Roozenbeek & van der Linden, 2018) showed a lower baseline resistance to misinformation, I speculated in this chapter that the intervention played a more prophylactic role for adolescents, where the intervention was likely more therapeutic in the present study. That is, adult participants in the present study are more likely to have been previously exposed to misinformation on the intervention topic, making this a therapeutic intervention for the sample. But the extent to which this is true, and individuals' alertness and resistance to and experience with misinformation is unknown. Findings from session-level quantitative variables also indicated that the level of interaction and score was not related to the size of the group. I suggested that individuals in larger groups may be showing signs of social loafing and intergroup anxiety, although this was highly speculatory. What is clear is that the internal mechanisms of the intervention (i.e., how players interacted with the intervention and each other) is of high importance, and further enquiry requires going beyond quantitative design.

The findings presented in this chapter are limited, as they only include results from quantitative measures. Although quantitative inquiry is essential in this research area, individuals' experiences with misinformation and information consumption is rich in content and context, which can only be considered through interacting with participants individually. As heralded in the comments above, a more comprehensive analysis of individuals' prior experience and how individuals interacted during the intervention is essential to better understand the effect of the intervention. Although

this chapter has summarised quantitative results from the present study, the study was mixed methods in design, including session observations and participant interviews. The next chapter in this thesis introduces rich contextualised insights to not only address the open questions raised here, but to further explore what kind of learning occurred during and after the intervention.

5. Multi-faceting Learning from an Analogue Game-based Misinformation Intervention

Parts of this chapter are taken from the paper by the author of this thesis 'Exploring the Learning Potential of a Game-based Intervention to Inoculate Against Misinformation', soon to be submitted to Social Media + Society.

Recent research into the longitudinal effects of game-based inoculation interventions have called for learning to be considered as part of the inoculation processes (Maertens et al., 2025). Yet, as highlighted in Chapter 2, existing models and theories on learning are yet to be drawn upon in this context. The exploration of learning as an intervention outcome requires the use of novel methods and data (see Chapter 3). Game-based inoculation research has remained largely quantitative, often adopting item evaluation tasks to measure intervention efficacy (Roozenbeek et al., 2024; cf. Cook et al., 2024). This has allowed scholars to evaluate the efficacy of interventions on large scales and at relatively low cost (Appel et al., 2025; Roozenbeek & van der Linden, 2020), but has neglected qualitative approaches that may help explain trends and behaviours. Following from the overview of quantitative results in the previous chapter, this chapter presents qualitative findings assessing the effectiveness of the game-based intervention *Fake News*. This chapter not only presents new findings but situates them in the context of the quantitative results previously discussed. This chapter aims to both contribute to RQ1 and answer the second research question:

> **RQ2.** What kinds of learning are stimulated by analogue game-based inoculation interventions against misinformation?

The findings presented and discussed in this chapter are drawn from the same study as the previous chapter: a within-subjects study assessing the efficacy of the game-based inoculation intervention *Fake News*. As well as the quantitative measures previously discussed, the second round of data collection included two qualitative elements, (a) session observations of 25 groups working collaboratively to complete the intervention, and (b) semi-structured interviews with 10 participants completed 15 to 18 weeks after the intervention (i.e., 6 to 9 weeks after the longitudinal post-test). Session observations were recorded in a field diary, and interviews were recorded, transcribed,

and coded using both descriptive and analytic codes to interrogate the data around themes of learning, misinformation, and game dynamics.

This chapter makes three contributions to the growing literature applying serious games to misinformation. First, this chapter uses a mixed methods research design to evaluate the intervention through the lens of learning theories such as experiential and collaborative learning. Qualitative methods also give deeper insights into participants' experiences and reflections on their involvement in the intervention, and how the intervention interacted with participants' existing knowledge and experiences of misinformation. The second contribution is a conceptual one, revealing how the careful design of game mechanics can facilitate different kinds of learning into the study of game-based inoculation interventions against misinformation. The final contribution is the consideration of the role inoculation theory plays in the design and outcomes observed from the study. In this chapter I argue that although participants had considerable prior working knowledge and awareness of misinformation (making the inoculation intervention therapeutic rather than prophylactic for many players), participants built social relationships over the course of their sessions which facilitated collaborative and experiential learning; the effects of which were not fully captured by the quantitative measures. I argue that these forms of learning may be closely attributable to the design of the intervention itself, suggesting that greater weight should be given to the design of interventions, both in discussion and practice.

This chapter is structured into four sections, considering different aspects of participant learning from the perspective of the player and the intervention's design. First, the chapter assesses and discusses how the in-person design of the intervention helped recollection and facilitated learning. Second, the chapter investigates the participants' experience of the intervention. Continuing from the previous chapter, the intervention's efficacy as an inoculation intervention is then assessed through a qualitative lens, focussed on evidence of learning. Finally, the chapter reflects on the design of the intervention, heralding effective elements of the intervention and highlighting areas where the intervention could be improved. The chapter concludes with a summary of these sections and a wider discussion of the design of game-based interventions.

5.1 Collaborative and Experiential Learning in Analogue Game-based Interventions

Research with game-based inoculation interventions such as *Fake News* have primarily relied on quantitative methods to measure specific outcomes, such as whether individuals are better able to label the veracity of items. However, as argued in the previous chapter, treating an intervention as a 'white box' (i.e., a process where the

internal workings of a system can be examined) enables a deeper understanding into why it has been (un)successful, and can guide the design of future interventions (Roozenbeek et al., 2024). Observations of groups completing sessions provide useful information on how individuals interact with the intervention and, in the case of this collaborative analogue intervention, learn from each other. This qualitative approach allows an interrogation of the intervention's efficacy through a pedagogical lens. In this section I overview how players learned from the intervention and from each other and discuss participants' impressions from the intervention.

The qualitative approach taken in this study allows deeper and richer insights into how participants learn from the intervention, allowing us to draw on the literature of pedagogy and learning theories. Experiential learning scholars theorise that knowledge is created through the transformation of experience (Kolb, 1984; Kolb & Fry, 1975), and scholars have tied serious games to this form of learning (Gee, 2003). A core feature of experiential learning (and the similar game-oriented 'probe, hypothesise, reprobe, rethink' cycle, see Gee, 2003) is reflection: the act of reflection allows an individual to create abstract conceptualisations and abstractions from an experience, constituting learning (Dewey, 1933; Mezirow, 1991; Mezirow, 1987). Donald Schön argued that two essential forms of reflection played a part in learning: reflection-in-action, and reflection-on-action (Schön, 1984). Findings from session observations and participant interviews demonstrate that many participants exhibited both.

Reflection-in-action can occur when, during an experience or intuitive act, an individual reflects critically on decision making in that moment, and challenging assumptions they make (i.e., on-the-spot learning). The intervention *Fake News* encouraged reflection-in-action in players, as individuals were forced to think critically about their ability to recognising manipulative techniques in text through the lens of a malicious character: "The group focus on Paula [for Political Posturing]'s methods and motivations, analysing the language used in each card. […] Both participants commented on how they found it very 'thought provoking'" (field diary, team 23). Moreover, this reflection-in-action was facilitated by cooperation between players, as hearing alternate viewpoints and perspectives encouraged individuals to reflect on their own decision making skills: "Most of the time I found it interesting just to see where other people are coming from" (Amber).

Reflection-on-action occurs as individuals reflect retrospectively on an experience, and qualitative feedback suggested that the intervention had, too, stimulated this. Participants' ability to recollect many details from the session after 15-18 weeks suggests that the intervention had made a sizable impression on participants, many of whom discussed having already spent time reflecting on the session before the interview (i.e., reflection-on-action): "but this definitely helped in the sense of now

when I do consume news I'm a bit more analytical about it and looking a lot more for: is this a reliable source? Are they just stating the facts or are they trying to frame this in a certain way?" (Alexander); "I might think about [what I'm reading] a bit more, because I think, oh, I did that game." (Wesley). During interviews participants also frequently reflected on how many of the misinformation techniques observed in the game are used in the real world, indicating the forming of generalisations and abstract concepts (another core part of experiential learning): "because the game itself makes you just think generally about the way it's been written and the way it might have been written in, say, a media outlet" (Julia). These recollections provide evidence that some participants have met some of the core stages of experiential learning.

Collaborative learning was also evident during sessions, the as players shared knowledge with one another through completing the intervention task. The intervention itself facilitated the development of social relationships between players, which began in a waiting area beforehand and continued throughout the session. Players communicated on key misinformation techniques related to the game: "[player 1] explains to [player 2] what clickbait is, with [player 3] also contributing to the conversation" (field diary, team 12). Players' discussions frequently went beyond the content in the game, applying misinformation techniques to real world news publications and politicians: "The group compared text on game cards to real life events and examples" (field diary, team 1); "Discussion is good and detailed. Comments are less about other [game] characters, but on what conspiracy actually is and what it contains" (field diary, team 16). This suggests that collaborative learning, and to a wider extent peer learning (i.e., where individuals learn through interactions with one another), potentially played a greater role in the intervention than expected. Scholars have also suggested that competition can positively influence collaboration and collaborative learning, providing a social motivation for 'productivity' (Deutsch, 1949; D. W. Johnson & Johnson, 1991; P. Williams & Sheridan, 2010). Competition (stimulated through the use of a communal leaderboard) was observed to increase engagement and motivation within sessions. The interplay between a competitive rewards system (between teams) and collaboration in this intervention is well summarised by Alexander: "I think it's just having that social element obviously turned into a bit of competition with the scores. That definitely made it feel a lot more fun […]. You know, I really enjoyed it. I think the group aspect really helped that process" (Alexander). The intervention itself, then, appears to have not only facilitated experiential learning, but also acted as a catalyst for collaborative learning between players as well.

A reoccurring theme when reflecting on the intervention was the conversations and interactions participants had with one another during play. Session observations carried out during the study gave an insight into the inter-personal interactions

between players, and how groups completed the intervention task together. With players working collaboratively, especially in larger groups of four, players were often required to compromise to reach a group consensus on in-game decisions. Some participants later reflected during interviews that this act of compromise had been difficult, especially between players with different backgrounds: “That bit was a little bit frustrating because we kept going over it” (Amber); “Frustrating. I found it frustrating that they couldn’t see the- especially the point about the ‘money changer’ [character]” (Robert); “Difficult, always difficult for me. I’m quite forthright about my opinions, but I don’t want to stifle people” (Curtis). In contrast, however, (often the same) participants said they had also found it “interesting” hearing the opinions of their fellow players, sharing interest in the conversations (often on polarising politicians and topical news, as recorded in the field diary) that had occurred during the session: “It was nice to discuss what we thought about different things, you know” (Dominic); “Most of the time I found it interesting just to see where other people are coming from” (Amber). Participants perceived the intervention as being valuable often due to these very interactions, and many suggested the intervention would be well suited to a school context: “I kept thinking about the possibility of using something very similar in a school setting” (Joan). It was highlighted that some of the value gained from the intervention was playing with other players they did not know or who had different social, political, or geographical backgrounds.

> *“But it is interesting to some extent like to know what older people than me like coming from. Like [from a] British background, like why they have such different point of view.” – Andrea*

Although the efficacy of the intervention *Fake News* as a platform for learning is encouraging, it may have little real-world impact if it has little uptake (Roozenbeek et al., 2024). A core theme that emerged in interviews was the appeal of the in-person nature of the study, and a clear sense that it had influenced participants’ decisions to take part. All participants reported a positive experience in the game session, using terms like “interesting” (all interviewed participants), “fun” (participants 1, 3, 7, 8, and 9), and “enjoyable” (participants 2, 6, and 10) unprompted. Wesley summed up the tone of many of the interviews, saying “I thought it was quite interesting because I don't get a chance to do anything like that particularly, and it makes you think in a different way”. Many felt that being able to work collaboratively with other players and the analogue in-person nature of the game had contributed positively to the experience. For example: “I think the group aspect made it a lot more fun” (Alexander) and “That [the intervention was a physical game] was enjoyable in its own right, and I like having to think, and I like meeting people. I like discussing things and working things out with other people” (Amber). The appeal of the in-person nature of the study was particularly

present in older interviewed participants, many of whom admitted being indifferent to the monetary incentive for participation: “Like I said, I don't even – you know – the Amazon voucher was a slight draw, but then, I don't even shop at it.” (Brooke); “We saw it and thought, well, that sounds quite interesting!” (Robert).

An important and often understudied (A. Johnson & Madsen, 2024; Roozenbeek et al., 2024) aspect of misinformation interventions is attractiveness. Members of the public are more likely to organically (and repeatedly) engage with an intervention that is more attractive and give value to the player (Davies et al., 2024; A. Johnson & Madsen, 2024), increasing its real-world impact. Players must be motivated to engage, and the in-person, collaborative nature of *Fake News* was an important aspect of this: “the whole experience for me, every bit of it I enjoyed. It's a game I’d quite like to do with my son” (Amber). Furthermore, some interview participants expressed that they would not have taken part if the intervention had taken place online: "you wouldn't have done it, would you, if it was online? (to Robert)" (Joan).

Recent research by Maertens et al. (2025) has highlighted the importance of memory for mediating resistance in game-based inoculation interventions. The adopted methodology in this study has allowed a unique view into how participants recollect certain aspects of such an intervention. It is especially noteworthy that, despite the interviews occurring 15-18 weeks after the game sessions, participants had strong recollections of the intervention and their experience. Tangible and collaborative elements of the intervention were best represented in recollections, with most participants able to remember the other players in their session, and to a slightly lesser extent, elements of the game itself, including characters, images, and game mechanics.

> *“There were four of us in the group. Me and another lady about my age, an old gentleman, and we had to- it was 4 different types of newspaper writing styles and we had to work out which belonged in which section” – Amber*

This section has presented and discussed qualitative evidence for learning and engagement from session observations and participant interviews. Despite the design of the intervention necessitating players to compromise to group decisions, individuals reported enjoying the process of collaborating with other players and sharing ideas. The intervention can also be viewed through a pedagogical lens. Collaborative elements of the intervention stimulated reflection-in-action, with individuals critically reflecting on their own skills in identifying misinformation techniques. More broadly, the intervention could be considered a collaborative learning exercise as individuals shared knowledge and understanding during play.

That the in-person, collaborative, tangible, and embodied nature of the intervention improved the intervention attractiveness aligns with recent research using

other interventions in in-person environments (Axelsson et al., 2024). It is surprising, then, that this in-person collaborative medium is poorly represented in literature (DeJong, 2023; Kiili et al., 2024) despite calls for a stronger focus on intervention attractiveness (A. Johnson & Madsen, 2024; Roozenbeek et al., 2024). Although more difficult to scale than digital interventions, in-person games appeal more to older individuals (Cross et al., 2023) who may be less digitally literate and therefore less willing to engage with digital interventions (cf. A. Johnson & Madsen, 2024). Secondly, although scholars have previously referred to game-based inoculation interventions as being 'experiential learning' (Appel et al., 2025; Roozenbeek & van der Linden, 2020; Saleh et al., 2024), scholars have not engaged with the mechanisms of these learning models (i.e., experiential, reflective, or collaborative learning). I find that how players engaged with the intervention and each other can be interpreted through these models. Given that scholars have recently called for misinformation-focussed inoculation theory to be considered more in terms of learning (Maertens et al., 2025), these models may help guide practitioners to design more effective interventions in the future.

5.2 Participants: Prior Knowledge, Motivation, and Methodology

Qualitative methods in particular can help to not only measure intervention efficacy, but contextualise findings based on individual participants themselves. This chapter presents analyses of findings through the perspective of participants, in particular a more comprehensive contextualisation of the prior knowledge of participants, their motivations for taking part in the research, and a novel reflection from participants on the methodology used in the study.

The prior knowledge of an individual has long been viewed as an important variable in the design of educational content (see 'Expertise Reversal Effect', Kalyuga et al., 2003). Educational psychologists posit that individuals with high knowledge in a topic learn by applying and updating existing cognitive schemas, a process different for individuals with little prior knowledge (J. R. Anderson, 1985; Bartlett, 1932; Kalyuga, 2007; Kalyuga et al., 2003). This is also true for inoculation theory. An inoculation intervention is most effective for an individual with a high enough prior knowledge to care about the subject, but not so high that they are already alert and aware of their attitudinal vulnerability on a topic. Results from the previous chapter found that, while the intervention was effective for adolescents in a previous study (Roozenbeek & van der Linden, 2018), adult participants in the present study did not exhibit significant improvement through quantitative measures. However, these measures were not able to ascertain participants' prior knowledge with the contextualised richness afforded by qualitative approaches. This section summarises and discusses qualitative findings from participant interviews on the prior knowledge and experience with

misinformation in participants. Additionally, this section further explores the motivations of participants to engage with the intervention, and individuals' reflections on the study's methodology.

Interview data suggests that participants had good prior knowledge of and experience with misinformation around news media, mirroring a trend of growing public attention and concern about this issue. The platforms and justifications voiced were diverse across participants but followed a trend of older adults consuming less online news from social media (e.g., Facebook) and instead relying on traditional print media through digital means (e.g., The Guardian, The Telegraph, Private Eye, and the Eastern Daily Press[13]) and television news channels (e.g., BBC News and Sky News). This is congruent with news consumption trends seen in the UK in 2024 (Ofcom, 2024a). All participants that were interviewed demonstrated a good awareness of misinformation and bias in news. Many interviewees cited their past experiences as a source for this heightened awareness. Common themes emerged in discussions, such as the effect of 'Brexit', political polarisation surrounding the Israeli–Palestinian conflict, and the COVID-19 pandemic. Some of these events, most commonly 'Brexit', was described as being eye-opening by participants.

> *"Brexit was a big thing that probably started making me really question... because you had so much that came out from like the Brexit campaign [that] was actually fake news and a lot of things that were being pushed, people believed. But (...) it was framed in a certain way to exaggerate the claims being made" – Alexander*

Participants also shared particularly personal experiences they had had of confronting misinformation: "We took my son around to stay with his Nan the other night, and his Nan tried to tell him that COVID didn't exist" (Amber). Furthermore, some participants explained they had felt conflict with family members by challenging their misinformed views on social media platforms. This suggests that many of the participants from the study had not only an understanding and awareness of misinformation, but that, for them, their relationship with misinformation was also personal.

> *"I had one or two family members – one cousin, one my older brother – who were kind of posting stuff about COVID and it was nonsense. It was paranoid nonsense. (...) I felt a kind of duty, I guess, to kind of challenge it. (...) in the end I actually fell out with my older brother and my cousin. We haven't really repaired the relationship since" – Curtis*

[13] The Eastern Daily Press is a local newspaper in the study's region.

Some participants also shared that they had in fact taken action to improve their 'news hygiene', including breaking out of their own perceived echo chamber: "So I've made, like, a more conscious effort to go out and make sure I know what other people think" (Amber). This understanding and awareness of misinformation is reflected by a recent government survey of voters in the 2024 UK general election, who cited misinformation and 'fake news' as one of their primary concerns (ONS, 2024).

Participants' prior knowledge and awareness of misinformation provide a possible explanation for the null results from quantitative measures presented in the previous chapter. Informed by original research with the same intervention (Roozenbeek & van der Linden, 2018), the quantitative measures used in this study are designed to measure the intervention's inoculating effect. Inoculation, however, is less effective when individuals have already been pre-exposed to the inoculating content (Compton et al., 2022). This is due to the intervention no longer being purely pre-emptive, and the lesser role played by mechanisms such as threat (Ivanov et al., 2022). Participants demonstrated a good understanding of their (and society's) vulnerability to misinformation in interviews and session observations, suggesting that the intervention may not have increased threat in participants: "I mean, I'm always expecting biased information" (Brooke). This is further corroborated by some participants describing the intervention as a "reminder" rather than feeling fresh awareness of their vulnerability to misinformation: "I think it serves as a reminder for me, especially when reading news and encountering figures and quotations" (Amber). This suggests that for many participants this intervention may have been more therapeutic in nature, a potential departure from the conceptual design of the game-based intervention.

Participants showed not only an existing knowledge and experience with misinformation but displayed interest in the study itself. When asked, many participants explained that, although the monetary reward for participation had been a slight incentive, curiosity had also been a motivator in taking part in the study: "Well, it sounded really interesting, to be honest" (Dominic); "It was partly for the money, but mostly because I'm curious and I just wanted to see what it was about" (Andrea); "Yeah, we do like a game! We do, we do like a game!" (Joan). During interviews, several participants also asked unprompted about the research study itself: "I mean I know I shouldn't be asking you questions, but did you find the younger age groups who had more difficulty in discussing things than the older broader age group?" (Dominic); "Team interested in the project and how the research is useful" (field diary, team 6).

Participants were also asked during interviews how they felt about the methodology of the study, and many participants were keen to give feedback on the use of 'fake news' testing articles and both Likert-scale measures and open-ended

questions asking individuals to list arguments they found persuasive/reliable and what they felt the writer was trying to convince them of. Assessing the effectiveness of interventions using inoculation theory as a foundational theory has been a challenging task for experimental psychologists, and an area requiring further research and exploration (Roozenbeek et al., 2024; but see Maertens et al., 2024). Responses from interviewees were mixed, further signalling the need for further investigation in this area. Some participants felt that the open-ended qualitative questions in particular captured their thoughts and feelings, and that it was a positive approach to capturing what they had learned during the session: "To me, it was obvious when I read the second one that I was thinking more, so I'm assuming it came over in my answers. And I know I wrote more!" (Amber); "I think more from the open-ended ones, because then also some of them you had to explain yourself. So that actually makes you pick out things and be like, you can't just put a number down" (Alexander). Despite this, some interviewees reported dissatisfaction with the testing approach as a whole, feeling that it didn't capture what they felt they had learned from the session.

> *"Yeah, I don't think it was enough that last bit to capture what the aim was of the whole thing. That could have been a lot different, [...], a little bit more making us think a bit more about it because it was a little bit – I don't want to say bland because it wasn't bland – it wasn't enough in-depth to capture how I felt, I think" – Dominic*

This section has summarised qualitative insights from session observations and participant interviews into the prior knowledge, motivation, and reflective views on study methodology from participants. Quantitative results from this study show that after completing the intervention, participants did not significantly view 'fake news' articles as being less persuasive or reliable, nor did they significantly agree with them less. Evidence from semi-structured interviews with a sample of participants suggests that this could primarily be due to a pre-existing heightened awareness and alertness to misinformation in participants, and that the design of the game *Fake News* makes incorrect prior assumptions on these players' ability (i.e., a different game-based intervention may have been better suited to the participants' 'skill level'). Despite this, as heralded in the previous section, participants did learn both experientially and collaboratively from the intervention. Perhaps, then, where the quantitative methodology was able to capture the inoculating effect experienced by adolescents in previous research, it was not nuanced enough to capture learning experienced by adult participants.

Assessing effectiveness of interventions using inoculation theory as a foundational theory has been a challenging task for scholars, and an area requiring further research and exploration (Roozenbeek et al., 2024). This is further confounded

by results from this chapter, which suggest that the learning gained from such an intervention is far broader than can be evaluated in item-evaluation tasks. However, participant interviews have given insight on this, with participants themselves feeling that their learning was better captured by open-ended survey questions and participant interviews. This may pave the way for further investigation into qualitative methods as approaches to evaluating the efficacy of inoculation interventions.

5.3 *Fake News*'s Efficacy as an Inoculation Intervention: Qualitative Evidence

Applied inoculation theory research has primarily focussed on intervention outcomes, and research that has measures inoculation theory mechanisms and novel concepts have included specially designed methodologies (Banas & Rains, 2010; Basol, 2022; Ma et al., 2023). These methodologies, such as Likert-scale survey questions and thought-listening exercises (Compton, 2012; Ivanov et al., 2013) have become standardised approaches to measuring inoculation outcomes. However, many inoculation mechanisms and moderating variables, such as counterarguing, threat, affect, and perceived involvement (i.e., prior knowledge of the issue) may be particularly suited to qualitative measures, with quantitative approaches often lacking contextual information or a 'full picture'. Within game-based inoculation interventions, few scholars have sought to identify key inoculation variables (Becker et al., 2025), and this research has often found null results (Ma et al., 2023). This section continues from the previous chapter on further analysing and discussing *Fake News*'s efficacy as an inoculation intervention. Specifically, this section presents evidence of post-inoculation talk, feedback exercises, and 'booster' interventions, and discusses the classification of this intervention as an inoculation theory intervention.

The inclusion of a feedback activity as part of the intervention proved valuable for players. Once players had finished the game, their created 'fake news' article was compared to the correct article for their character and feedback was provided where the group had chosen incorrectly[14]. Session observations recall this part of interventions being particularly engaging, with participants often being keen to understand and discuss mistakes made: "The group score 7[/9], but are keen to understand where they have made mistakes" (field diary, team 7); "The team finished with time to spare, but were eager to know where they went wrong as they only scored 3[/9]" (field diary, team 16). Recent research has highlighted the importance of feedback in game-based interventions (Leder et al., 2024; early work by Pearson & Smith, 1985), suggesting that such exercises can strengthen conferred resistance and improve

[14] This marks a deviation from original research with the intervention *Fake News*, which gave feedback to players after the post-test.

memory retention. Evidence from interviews supported this, with the team's score and feedback activity being well recollected by participants. Moreover, the feedback activity provided an opportunity for participants to reflect on their decision-making, stimulating reflection-on-action (Schön, 1984).

The intervention may have also acted as a catalyst for participants to share knowledge outside of the intervention, raising awareness within their social network. Inoculation is often seen as a subvocal process; however, post-inoculation talk (PIT, i.e., word-of-mouth communication post-intervention) may be a platform for positive effects to transfer between individuals as they look to 'advocate their position with others' (Basol, 2022; Compton, 2024; Compton & Pfau, 2009). One participant described having PIT within their social network of friends immediately after the intervention.

> *"I talked to people afterwards as well – like straight afterwards. 'Ohh I did this'; 'This was really interesting'; 'We did this and' – just because I just found it fascinating, and like I said, the dynamics between the four of us working it out together and how it worked" – Amber*

Although it is unclear if Amber shared arguments or skills in conversations with peers, this demonstrates the PIT potential for this type of intervention. It is also of interest as the quote suggests that PIT was stimulated through the novel design of the intervention (i.e., a collaborative analogue card game). Games may be a particularly good catalyst for PIT as their relatable and informal nature may encourage discussion within social networks, and for public interventions, may promote organic uptake.

There is also evidence that the external stimulus (i.e., the longitudinal post-test) had a positive impact on participants, acting as a memory-strengthening 'booster' activity. Many interviewees found the follow-up survey was able to "serve as a reminder" (Amber) of both prior knowledge and learned techniques, highlighting existing concerns of testing effects (Roozenbeek et al., 2024), and demonstrating the positive effects of booster interventions: "Yeah, it definitely was a good way to test it because you're getting someone to read a realistic news article and then apply sort of what they [have learned]" (Alexander); "Yeah, it jogs it back into place in your memory. It makes you think again." (Amber). Such testing effects have been discussed in previous research (Capewell et al., 2024; Maertens et al., 2021; Roozenbeek et al., 2020), but few qualitative studies have found participants to be self-aware of a stimuli's memory-strengthening benefit. The findings further support the potential for "booster" activities to strengthen intervention outcomes, particularly where this activity applies techniques learned in the original intervention: "it definitely was a good way to test it because

you're getting someone to read a realistic news article and then apply [...] what they've learned" (Alexander).

It is important to recognise that many of the observed outcomes of this intervention can be attributed to specific design decisions of the intervention. For example, as shown earlier in this chapter, the collaborative analogue genre of the game *Fake News* had a direct impact on stimulating word-of-mouth communication in players after the intervention, potentially spreading the learning outcomes beyond the study. It also facilitated the development of social relationships between players during the intervention, providing a platform for collaborative and experiential learning. Moreover, tangible, visual elements within the design of the game positively impacted memory recall of interviewees, and qualitative findings also support the replayability potential of analogue serious games. Yet, design decisions and design processes that have led to these outcomes are rarely discussed in literature. That the design of *Fake News* was so impactful to the outcomes of the intervention suggests that the same may be the case for other misinformation interventions, and we argue that this deserves greater focus in game-based misinformation research (see Leder et al., 2024; Roozenbeek et al., 2024). Further consideration of game design in this intervention may have, for example, led to adaptations of content more appropriate for the prior knowledge of these participants, an observation made in the previous section.

The importance of considered design brings into question the role that inoculation theory plays in game-based misinformation interventions. Conventional inoculation interventions can elicit threat and improve counterarguing skills through broadly agreed upon message components: an explicit forewarning and refutational pre-emption tasks. How inoculation mechanisms can be stimulated in game-based contexts is less clear, and how this is done is largely down to the interpretation of the practitioner designing the intervention. The difficulty of designing interventions to stimulate inoculation mechanisms is further complicated in therapeutic technique-based interventions, where message content broadly delivers educational material (instead of pre-emptively refuting specific arguments on a topic), and the role of threat itself is less clear (Ma et al., 2023; Maertens et al., 2025). This has been the case for the participants taking part in this intervention, many of whom demonstrated preexisting awareness and alertness to misinformation through world events and personal experiences. This raises the question of whether, for technique- and game-based interventions such as *Fake News*, it is still appropriate to consider inoculation as being the core foundational theory, or whether a greater focus on serious game pedagogy and memory creation may better guide practitioners in creating more effective interventions in these contexts (Maertens et al., 2025). Game-based inoculation interventions such as *Fake News* sit at the intersection between serious games and

inoculation theory, and further discussion on and consideration of design approaches are essential to creating more effective interventions.

This section has summarised the game-based intervention *Fake News* from a qualitative perspective, presenting evidence for popular inoculation theory concepts such as post-inoculation talk, feedback exercises, and 'booster' activities. Moreover, this section stresses the importance of design processes and design decisions in facilitating a game-based intervention to achieve outcomes. I argue that, while at present much attention is given to inoculation theory as a conceptual foundation to the intervention, this intervention is in reality much detached from inoculation theory mechanisms and components. Giving more attention to serious game design and serious game pedagogy may help guide practitioners to create more effective interventions.

5.4 Considering *Fake News* Through the Lens of Game Design

The previous section has argued that greater attention should be paid to serious game design processes and decisions in game-based inoculation literature to create interventions with research-supported design choices. Scholars have recently called for further interrogation of the intervention design process (Roozenbeek et al., 2024), and recent research by Hopkins et al. (2023) has found success holding co-design workshops with key stakeholders to adapt the intervention *Cranky Uncle Vaccine* to different cultural contexts. The present study took an existing game 'off-the-shelf', drawing from another influential intervention by Roozenbeek and van der Linden (2018), rather than creating a new game from scratch. However, pedagogical and serious game design literatures can be used to retrospectively assess the design of the intervention used in the study. This section seeks to evaluate the design effectiveness of the intervention through the lens of (serious) game design literature and memory research, finding that although the intervention was robust to memory fade, and facilitated humorous interactions and different types of learning (see Section 5.1), there is room for improvement in design of the intervention to better facilitate player collaboration and broader types of 'transformations'.

Despite the lack of engagement with game design literature from game-based inoculation scholars, research has explored how humorous design can improve intervention effectiveness (Compton, 2018; Vraga et al., 2019). Specifically, scholars posit that humour can improve post-inoculation talk, improve memory retention, improve the entertainment value of the intervention, and make serious topics more accessible (Hopkins et al., 2023; Moyer-Gusé et al., 2018; Roozenbeek & van der Linden, 2020). In particular, humorous design is posited to reduce reactance of individuals during inoculation interventions on sensitive or polarising topics, especially where

individuals may have opposing prior views (Hopkins et al., 2023; Moyer-Gusé et al., 2018; Roozenbeek & van der Linden, 2020; Vraga et al., 2019). In the present study, players collaboratively represented a fictitious character in a fictitious setting (albeit within the very polarising issue of immigration). Character cards featured humorous writing, as did the rules briefing directly before the intervention. Observations during sessions found that, despite detailed discussion and comparisons between characters and real-world politicians, individuals frequently joked about characters and the task: "Light laughter and humour observed in this group" (field diary, team 6); "The team joke as they complete the intervention" (field diary, team 8); "[redacted] jokes that Mike the Money Changer is a 'lovely chap'" (field diary, team 12); "[Players were] laughing when reading and looking at some of the cards" (field diary, team 17). Moreover, asking players to play from the point of view of the character (rather than themselves) may have allowed players to discuss more freely. Indeed, over all 45 sessions from both rounds of data collection, no tension was recorded to have arisen from political differences between participants.

Other design elements of the game-based intervention *Fake News* were also well recollected by participants during interviews, supporting their memory-strengthening potential. For example, the names of some characters were recalled on multiple occasions unprompted: "I remember one because it was Dana and Fox and I'm a great fan of the X Files!" (Curtis). Some character names made reference to fictional characters in popular culture, which may have helped with memory retention. Despite the image cards (see for example, Figure 21) making up only small portion of game content, these were also disproportionally well recollected by participants over other textual game cards: "we try to match […] the images to align with the characteristics and their political point of view. That's what I remember, and I remember doing it in teamwork" (Andrea). This superiority of imagery over text in memory retention is congruent with literature on memory and cognition (see Picture Superiority Effect, Shepard, 1967; Standing, 1973). The resilience of popular culture references and imagery in game content to memory fade demonstrates their potential for reducing inoculation decay in future interventions. Understanding the potential role of memory characteristics such as the picture superiority effect (Shepard, 1967) and the multimedia effect (Mayer, 2002) in serious games may help practitioners design effective interventions with more robust memory retention.

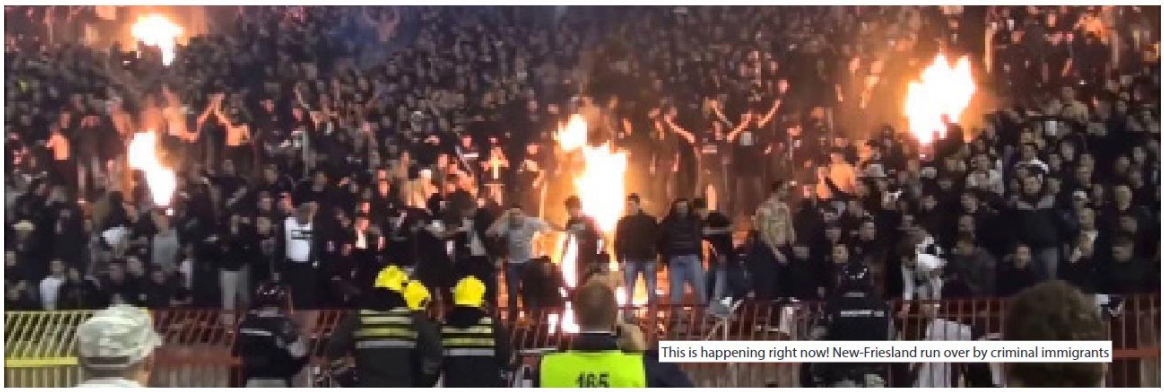


Figure 21: An image card for the intervention character 'Mike the Money Changer'.

In this section I aim to not only highlight positive outcomes from elements of the intervention but also identify areas where the design of the intervention could be improved. The section assesses the quality of the intervention as both a collaborative analogue game and a serious game, drawing on work by Zagal et al. (2006) and Schell (2019). The intervention *Fake News* asks players to work collaboratively to write a 'fake news' article in the style of a given character. Effective collaborative design is difficult to achieve, however, as designers must not only create a game that is engaging and enjoyable, but that elicits small-group interactions through which players effectively collaborate to achieve a shared goal (Peppler et al., 2013). Zagal et al. (2006) identified three core challenges associated with designing effective collaborative games, which act as useful markers when assessing how effectively the intervention invites collaboration in players.

One challenge of collaborative gameplay design is effectively providing rationale for collaboration, to avoid "one player making all the decisions for the team" (Zagal et al., 2006, p. 32). This is a key challenge to overcome as, were one player to dominate discussion and decision making for the team, reflective and collaborative learning would be diminished (i.e., players may feel less able to share contrary viewpoints and share key knowledge relevant to the task), and players may enjoy the intervention less. Observations made during sessions found that some players were more dominant during play: "player 1 takes the lead, but discourse is shared" (field diary, team 3); "player 1 does lead the session a little, but isn't intrusive or rude" (field diary, team 9). Furthermore, session observations recorded how some players took such a dominant role in the group that other players were not able to participate: "player 3 does not really get much of a word in, as player 1 and player 2 lead the session by the end" (field diary, team 9); "player 2 is often wanting to say things but is not able to speak given player 1's dominance and volume" (field diary, team 13). The previous chapter explored how the level of interaction did not increase with larger groups and that this was perhaps due to social loafing and intergroup anxiety. These findings

further support this, as dominant players in groups may have also increased social loafing and intergroup anxiety.

Ensuring participants are interested in the outcome of the game is also important for not only collaborative engagement but replay as well (Zagal et al., 2006). In the previous chapter of this thesis, I considered whether the reason for groups not interacting more when playing as harder characters was due to a poor feedback loop in the intervention's design. Put simply, the game does not offer any form of feedback to players to inform them if they are completing the task (un)successfully, and their score is only communicated to them at the end of the intervention. Similarly, the lack of any form of feedback loop built into the game means there is little sense of jeopardy during play. Session observations record groups finding the end of the game satisfying and engaging: "[The group are] very engaged and invested when the answers were read out" (field diary, team 18). However, this has little motivating power during play unless the individual has already played before, being aware of the positive affective response when achieving a good score. Intervention uptake has been repeatedly identified as a critical challenge (A. Johnson & Madsen, 2024; Roozenbeek et al., 2024), but maintaining engagement with an intervention after play has begun also requires attention.

A third challenge of collaborative game design is it must deliver different experiences each time a player engages with it, otherwise players may only engage with it once (Zagal et al., 2006). As highlighted in the comments above, intervention engagement and replayability are pressing concerns in inoculation research (A. Johnson & Madsen, 2024; Maertens et al., 2025; Roozenbeek et al., 2024). Finding ways to encourage individuals to engage with an intervention is essential as longer-lasting resistance has been best achieved through "booster shots" (Maertens et al., 2025; but see Ma et al., 2023). The intervention *Fake News* asks groups to play as one of four characters, which they assume throughout the session. To this end, the intervention *Fake News* does support four different playing experiences, and session observations found that players expressed an interest in playing as the other characters in future: "All three players play standing around the table, and player 1 remarks that it would be fun to sort all characters" (field diary, team 4).

The serious game design literature inherits many concepts and theories from pedagogy and learning theory. For example, I have previously discussed the importance of learner prior knowledge when designing educational content. This is also a central concept in serious game design literature (Aleven et al., 2010; Annetta, 2010; Culyba, 2018; Schell, 2019), and poor consideration of this concept can also affect motivation and engagement with an intervention. If the intervention content is too simple the player can become disengaged and bored. Too hard, and the player can

become frustrated (Annetta, 2010). Findings from the present study suggest that the prior knowledge of and awareness of the participants was higher than anticipated (see Section 4.2), with the intervention topic being a relevant issue in British politics, and individuals having shared personal experiences with misinformation. A more challenging task may have led to more effective learning (see 'Expertise Reversal Effect', Kalyuga et al., 2003), but an intervention that is more engaging to all players.

Serious game design scholars have also identified a broad spectrum of learning (i.e., player transformations) that individuals may get from a serious game and have highlighted that this is often far wider than the measures used to establish serious game effectiveness (Culyba, 2018). The Transformational Framework[15] (Culyba, 2018), a practitioner-oriented serious game design framework (see Chapter 3), suggests that serious game-based learning goes beyond knowledge and skill, and can change players' behaviour, beliefs, disposition, relationships within society, and even physical health (Culyba, 2018). The quantitative measures adopted in the present study (i.e., asking participants to rate the persuasiveness, reliability, how much they agree with 'fake news' articles, and identify unpersuasive/unreliable arguments in these articles) measure participant behaviour towards articles written with misleading writing techniques. As heralded in the sections above, however, players also experienced 'transformations' in knowledge (e.g., learning misinformation techniques from each other) and experience (the intervention forms an experience on which players can later reflect). The design of the intervention could also be improved to facilitate other types of outcomes from players, such as improving misinformation identification skills (a skill that is called upon in the intervention but not trained or developed).

This section has interpreted qualitative findings from the present study through the lens of serious game design and pedagogy. Specifically, I have presented evidence on the ability of humorous design to reduce player reactance in divisive topics, and the ability for multimedia elements to strengthen memory retention of the intervention. This section also assesses the effectiveness of collaborative design elements in the intervention (Zagal et al., 2006), finding that the intervention does not provide enough rationale for all players to participate equally and lacks effective feedback loops to motivate players during play. Finally, this section assesses the theoretical effectiveness of the intervention with reference to practitioner-oriented serious game design frameworks. I find that better consideration of players' prior knowledge in intervention design could have improved its effectiveness, and a broader scope of learning may better explain findings from previous sections in this chapter.

[15] Games designed to serve an educational purpose were labelled as 'serious games' (and are done so in this thesis) in the 1970s (Abt, 1970), but many scholars have since proposed more descriptive names such as 'transformational games', 'applied games', and 'educational games'.

Summary and Conclusion

This chapter has demonstrated how peoples' experiences of and engagements with misinformation are highly contextual. Humans do not take on new information as simply fact, but consider its veracity and reliability based on the context in which it was read (Pennycook & Rand, 2019; Traberg & van der Linden, 2022), how it relates to the prior beliefs of the person, and a broad spectrum of further biases (Bolsen et al., 2014; Brady et al., 2020; Ecker et al., 2022; Roozenbeek & van der Linden, 2024). Conventional inoculation theory research has in the past used qualitative methods such as thought listing or think aloud exercises to measure the effectiveness of inoculation interventions (Ivanov et al., 2013). Despite this, game-based inoculation research has relied on quantitative methods as they give scholars the ability to easily scale-up interventions that be carried out quickly, and qualitative methods have not yet been fully explored in this context. This means that quantitative measures of intervention effectiveness do not tell the whole story. This chapter has sought to present qualitative findings to answer the second research question of the thesis: what kinds of learning occur within analogue game-based inoculation interventions against misinformation.

First, I have reviewed how individuals felt about the intervention and working with others, finding that the tangible, embodied, and collaborative nature of the intervention facilitated collaborative and experiential learning. Participants learned from one another during play and exhibited evidence of reflection-in-action and reflection-on-action on the experience. Secondly, qualitative interviews give an insight into the participants themselves, finding that participants already had good prior awareness and alertness to misinformation. Thirdly, the efficacy of the intervention is assessed through the lens of inoculation theory, finding evidence of post-inoculation talk, and the efficacy of 'booster' interventions. I also interrogate inoculation's role in the intervention *Fake News*, questioning inoculation as the intervention's foundational theory. Finally, this chapter assesses the intervention's design. Making reference to both collaborative and serious game design literature, I identify a number of shortcomings in the design of *Fake News* and discuss the potential contribution of practitioner-oriented serious game design frameworks.

This chapter offers three contributions to game-based inoculation literature. First, qualitative results from this chapter highlight the potential for both analogue and collaborative intervention design, an important finding as game-based interventions are most commonly digital (DeJong, 2023; Kiili et al., 2024). Session observations detailed how the intervention's tangible, embodied, and analogue design allowed players to build social relationships over the course of their session, creating a memorable and positive experience for players, and providing a platform for collaborative learning as players completed the intervention together sharing

knowledge during play. Evidence suggested that the intervention also facilitated experiential learning, with recollections from participants in interviews highlighting the intervention's concrete nature as an experience, evidencing bot reflection-in-action and reflection-on-action, and demonstrating abstract comparisons to real world contexts.

The second contribution of this chapter is a demonstration of the potential for qualitative methods in game-based inoculation research. Although quantitative inquiry is essential in this research area (i.e., to measure empirical validity of an intervention), qualitative methods play an important part in providing rich contextual insights into the intervention, and the players themselves. In the previous chapter, quantitative measures showed no significant improvement in persuasiveness, reliability, and personal agreement judgements of 'fake news' testing articles after taking part in the intervention. Qualitative findings, however, demonstrated how these null results may be due to existing prior knowledge of participants, and the learning from the intervention was in fact broader than quantitative measures. Moreover, session observations gave key insights into how players communicated with one another during play and interacted with the intervention. Qualitative findings presented in this chapter not only give a broader and deeper insight into the effect of the intervention but can also guide practitioners to create more effective interventions in the future.

The final contribution of this chapter is a retrospective review of inoculation theory's role in the design of the intervention *Fake News*, and a consideration of serious game design literature to both help design and interpret findings. Although evidence from interviews and session observations supported contemporary ideas within inoculation theory (i.e., post-inoculation talk and 'booster' session), the alertness and prior knowledge of misinformation demonstrated by participants indicated that parts of the intervention played a more therapeutic role (rather than acting pre-emptively). We argue that the role of inoculation in this therapeutic intervention is particularly unclear, and many of the outcomes from the intervention could be more closely attributed to the design of the game itself, rather than increases of awareness or alertness. Moreover, prominent ideas from serious game design literature such as the Transformational Framework (Culyba, 2018) aligned with the findings from this study, suggesting that this literature can not only help in the design of effective interventions, but in the interpretation of results.

This study is an encouraging demonstration of how mixed-methods approaches can be used in game-based inoculation research, allowing for greater insights into individuals' prior knowledge, the interaction between players and the serious game, and how the intervention conferred resistance to players. Furthermore, we interpret these findings as a call for deeper investigation and experimentation into

how serious game design decisions affect intervention effectiveness, and the role that inoculation theory has in this design process. Future research could look at both how different mechanics confer resistance differently (e.g., analogue vs. digital interventions), or how different design processes (such as co-design workshops, see Cook et al., 2024) could lead to more effective interventions that perform better in real-world environments.

Inoculation theory has been an accepted foundational theory for recent game-based interventions against misinformation. Despite inoculation theory having informed the design of the intervention *Fake News*, findings have struggled to identify inoculation mechanisms being stimulated. The intervention failed to elicit new threat in players, as interviews demonstrated that many participants already had a heightened sense of awareness and alertness to misinformation. Many participants also reported the intervention acting more as a "reminder" than learning new misinformation techniques or counterarguments. Moreover, quantitative and qualitative results found that the intervention did not increase negative affect in players, but in some cases increased positive affect. Despite this, the intervention was successful in eliciting collaborative and experiential learning wider than the content of the intervention itself. Findings also indicated that interaction with collaborative and serious game design literature may help design more effective interventions and interpret results. The findings from this chapter not only raise questions of inoculation theory's role in game-based interventions but assesses the role that serious game design literature has to offer in this context. The next chapter of this thesis seeks to answer these questions, comprehensively reviewing how scholars have interpreted the mechanisms of inoculation theory in the design of game-based interventions, and how deeper discussion and exploration of intervention design processes and decisions may help to create more effective interventions.

6. Inoculation Theory as a Design Approach to Game-based Misinformation Interventions: A Review

Parts of this chapter are taken from a publication by the author of this thesis, 'Inoculation Theory as a Design Approach to Game-Based Misinformation Interventions: A Review', in the journal Popular Communications.

Game-based inoculation research has been applied to important global issues such as climate change misinformation (Cook, 2021), extremism (Saleh et al., 2023), and vaccine hesitancy (Cook et al., 2025). Despite this, relatively little scholarship exists investigating the game components and internal mechanisms that improve or worsen intervention effectiveness (Appel et al., 2025). This is in contrast to conventional research into inoculation theory, where mechanisms such as threat and counterarguing (Banas & Richards, 2017; Compton, 2021; Ivanov et al., 2022), moderators such as affect and issue familiarity (Compton et al., 2022; Pfau et al., 2001), and message components such as forewarnings and refutational pre-emption tasks have seen much interrogation (Lewandowsky et al., 2012; Parker et al., 2025; van der Linden, 2022). Scholars have more recently begun to investigate game-based intervention design more closely (see Appel et al., 2025; Leder et al., 2024; Maertens et al., 2025), but this research is limited and has not yet interrogated the role that inoculation plays in achieving outcomes. Findings from the previous chapter have questioned the role inoculation theory plays in the game-based intervention *Fake News* and call for a deeper assessment of the theory's use in game-based contexts. This chapter aims to achieve this, and addresses research question 3:

> **RQ3**. How could game-based inoculation interventions against misinformation be better designed?

This chapter presents findings from a narrative integrative review of 25 misinformation serious games that use inoculation as a foundational theory. 25 games and 22 corresponding papers were identified through searches on Google Scholar and through recursive searches of references in related publications. Playtesting of serious

games took place between September and December in 2024, and serious games have been reviewed from both a game design and inoculation theory perspective. The variables used to review the collated games and research articles have been selected to review a) the functionality and gameplay of each serious game, b) the discussion and design around design approaches adopted, and to c) facilitate discussion on the interpretation and implementation surrounding inoculation theory in game-based contexts. The game-based interventions included in this review are listed in Table 7. Papers in which the reviewed interventions are first presented are also listed, as this informs discussion later in this chapter.

*Table 7: Details of the game-based interventions included in the review, and the papers in which these interventions are first introduced. * indicated on intervention designer when none are explicitly specified in paper.*

Intervention name	Paper Id	Authors	Year	Intervention designer
Fake News	1	Roozenbeek, J., & van der Linden, S.	2018	Authors and DROG
Bad News	2	Roozenbeek, J., & van der Linden, S.	2019	Authors, DROG, Gusmanson, and TILT
Harmony Square	3	Roozenbeek, J., & van der Linden, S.	2020	Authors, DROG, Gusmanson, and TILT
ChamberBreaker	4	Jeon, Y., Kim, B., Xiong, A., Lee, D., & Han, K.	2021	Authors
Go Viral!	5	Basol, M., Roozenbeek, J., Berriche, M., Uenal, F., McClanahan, W. P., & Linden, S. van der.	2021	Authors, DROG, Gusmanson, TILT, and UK Cabinet Office
Cranky Uncle	6	Cook, J.	2021	Author and Goodbeast
Join This Group	7	Harjani, T., Basol, M.-S., Roozenbeek, J., & van der Linden, S.	2023	Authors and Digital Empowerment Foundation (DEF)
Fake News Detective	8	Tandoc, E., & Seet, S.	2023	NTU School of Computer Science and Engineering
Cranky Uncle Vaccine (East Africa edition)	9	Hopkins, K. L., Lepage, C., Cook, W., Thomson, A., Abeyesekera, S., Knobler, S., Boehman, N., Thompson, B., Waiswa, P., Ssanyu, J. N., Kabwijamu, L., Wamalwa, B., Aura, C., Rukundo, J. C., & Cook, J.	2023	Authors and Goodbeast
Misinformation Is Contagious	10	Barzilai, S., Mor-Hagani, S., Abed, F., Tal-Savir, D., Goldik, N., Talmon, I., & Davidow, O.	2023	Authors

Follow Me	11	Pomichal, V., & Trnka, A.	2022	Impact Games and Doublequote Studio
Spot The Troll	12	Lees, J., Banas, J. A., Linvill, D., Meirick, P. C., & Warren, P.	2023	Authors and Media Forensics Hub Clemson University
COVID-19 Fake News (CFN)	13	Rędzio, A. M., Izydorczak, K., Muniak, P., Kulesza, W., & Doliński, D.	2023	Authors*
Unnamed 1	14	Ma, J., Chen, Y., Zhu, H., & Gan, Y.	2023	Authors
Radicalise (Iraq edition)	15	Saleh, N. F., Makki, F., van der Linden, S., & Roozenbeek, J.	2023	Authors
Werewolf On Campus	16	Adams, A. D.	2023	Authors
Cat Park	17	Neylan, J., Biddlestone, M., Roozenbeek, J., & van der Linden, S.	2023	Authors, Gusmanson, and TILT
Covid-19 Trivia, *Myth Busters*, and *Unnamed 2*	18	Grace, L., Dunleavy, V. O., Ahn, R., & Mayo, D.	2024	Authors
Radicalise	19	Saleh, N. F., Roozenbeek, J. O. N., Makki, F. A., Mcclanahan, W. P., & van der Linden, S.	2024	Authors
Cranky Uncle Vaccine (Ghana edition)	20	Cook, J., Lepage, C., Hopkins, K. L., Cook, W., Kolog, E. A., Thomson, A., Iddrisu, I., & Burnette, S.	2024	Authors and Goodbeast
Unnamed 3	21	Orticio, E., Meyer, M., & Kidd, C.	2024	Authors
Bad Vaxx	22	Appel, R. E., Roozenbeek, J., Rayburn-Reeves, R., Basol, M., Corbin, J., Compton, J., & van der Linden, S.	2025	Authors, DROG, Gusmanson and TILT

This chapter makes three contributions to the game-based inoculation literature. Firstly, this review finds a lack of exploration of different types of game-based intervention. Specifically, the review finds that the potential affordances of different platforms (e.g., analogue design or virtual reality) have been underexplored in place of digital interventions that can be more easily scaled. Furthermore, I find that many digital game-based interventions follow a similar design to an early and high-profile intervention called *Bad News*, with little exploration of whether its design can achieve the same effect with different audiences and new applications, or whether other designs may be more effective. Secondly, this review finds that discussion and transparency around how inoculation has been integrated into the design of game-based interventions is often superficial or lacking. Very rarely is inoculation theory

framed in discussion as a driving concept informing the design of the intervention. Finally, this review finds little discussion on design processes adopted, a key part of serious game design and something that serious game design literature highlights as a key contributor of intervention effectiveness.

This chapter is structured into four sections. The first section addresses the findings from the review relating to game design, presenting results on a number of variables including game topic and genre, target audience, and platform. The second section presents evidence of how inoculation has been realised in the reviewed game-based interventions, including key mechanisms, components, and related concepts such as feedback exercises. The third chapter looks beyond how the reviewed interventions elicit inoculation theory mechanisms and takes a more analytical tone, interrogating inoculation theory's role as a driving design concept. The final section of this chapter reviews how reviewed games' corresponding papers discusses design processes and design considerations. This chapter argues that serious game design literature may offer more instructional support to practitioners designing new interventions, and further exploration of (serious) game design frameworks may yield more effective interventions.

6.1 Overview of Current Design Trends in Game-based Misinformation Interventions

Although other scholars have previously reviewed the design of game-based inoculation interventions (see DeJong, 2023; Kiili et al., 2024), these have included serious games that do not draw on inoculation theory. These reviews highlight trends in the design of serious games of misinformation, but as many of these serious games draw on other communicative theories (e.g., transportation theory, see Grace & Liang, 2023), these findings are not generalisable to game-based inoculation theory interventions as a whole. This section presents findings on game design from the review of game-based misinformation interventions. Specifically, this section reviews the serious game topic, genre, platform, target audience, and roleplaying perspective. This section also offers analysis of findings, discussing the consequences of the under exploration of platform and genre for the field as a whole. Table 8 summarises the review's findings on these game design variables.

Table 8: Game design variables of game-based intervention. Findings compiled from corresponding paper and, where possible, playtesting completed by the author. N.A. indicated on roleplaying perspective where intervention does not involve roleplaying.

Intervention name	Topic	Genre	Platform	Target audience	Roleplaying perspective
Fake News	Misinformation on immigration	Collaborative phrasal template card game	Analogue	Initial research with Dutch high school students	Bad
Bad News	Misinformation	Linear roleplaying narrative text adventure	Browser	General public (14+)	Bad
Harmony Square	Political misinformation	Linear roleplaying narrative text adventure	Browser	General public (15+)	Bad
ChamberBreaker	Misinformation echo chambers	Linear roleplaying narrative text adventure	Browser	General public	Bad
Go Viral!	Vaccine misinformation	Linear roleplaying narrative text adventure	Browser	General public (15+)	Bad
Cranky Uncle	Climate misinformation	Menu-driven linear quiz narrative	Browser & mobile	Initial research primarily with U.S. university students	N.A.
Join This Group	Messaging application misinformation	Linear roleplaying narrative text adventure	Browser	General public in India	Good
Fake News Detective	Misinformation	Narrative roleplay adventure	Browser	Young players, aged 16 and above	Good
Cranky Uncle Vaccine (East Africa edition)	Vaccine misinformation	Menu-driven linear quiz narrative	Browser & mobile	General public in Uganda, Kenya, and Rwanda	N.A.

Misinformation Is Contagious	Misinformation sharing	Linear roleplaying narrative text adventure	Browser	Initial research with Israeli middle school students	Good
Follow Me	Misinformation	Strategy Simulation Pattern Recognition Card Game	Analogue	High school students	Good
Spot The Troll	Bot/Troll identification	Quiz	Browser	General public	N.A.
COVID-19 Fake News (CFN)	Vaccine misinformation	Linear roleplaying narrative text adventure	Analogue	General public	Bad
Unnamed 1	Vaccine misinformation	Quiz	Browser	Initial research with Chinese adults between 18 and 55	N.A.
Radicalise (Iraq edition)	Radicalisation	Linear roleplaying narrative text adventure	Browser	Initial research with Iraqi adults	Bad
Werewolf On Campus	Misinformation	Collaborative narrative roleplay	Analogue	Initial research with university students	Good
Cat Park	Multimodal misinformation	Narrative visual novel	Browser	General public (15+)	Both
Covid-19 Trivia	Health misinformation	Quiz microgame	Browser	General public	N.A.
Myth busters	Health misinformation	Roleplay quiz clicker	Browser	General public	Good
Unnamed 2	Health misinformation	Roleplay fictional narrative microgame	Browser	General public	Good
Radicalise	Radicalisation	Linear roleplaying narrative text adventure	Browser	Initial research with UK adults	Bad
Cranky Uncle Vaccine (Ghana edition)	Vaccine misinformation	Menu-driven linear quiz narrative	Browser & mobile	Initial research primarily with U.S. university students	N.A.

Unnamed 3	Novel claims	Quiz clicker microgame	PsychoPy	Initial research with young children between 4 and 7	N.A.
Bad Vaxx (good edition)	Vaccine misinformation	Linear roleplaying narrative text adventure	Browser	General public	Good
Bad Vaxx (evil edition)	Vaccine misinformation	Linear roleplaying narrative text adventure	Browser	General public	Bad

Game Topic

I have categorised the reviewed game-based misinformation interventions as being in one of two topic groups: information topics, and misinformation techniques. Games such as *Bad News*, *ChamberBreaker*, *Misinformation is Contagious*, *Spot The Troll*, *Follow Me*, *Join This Group*, *Fake News Detective*, *Werewolf On Campus*, and *Cat Park* inoculate players on misinformation phenomenon (e.g., echo chambers, misinformation sharing, and misinformation in direct messaging), or techniques used in the misinformation creation or dissemination process. Information topics currently covered by interventions include misinformation on immigration (*Fake News*), politics (*Harmony Square*), radicalisation/extremism (*Radicalise*), climate change (*Cranky Uncle*), and novel claims (*Unnamed 3*). The growth of game-based inoculation interventions coincided with the COVID-19 pandemic and has resulted in a disproportionally greater attention towards vaccine and health misinformation: eight games focus on health misinformation and vaccine hesitancy. It is important to note that games inoculating against information topics often do so by inoculating against techniques often used in those topics. For example, *Radicalise* inoculates against the techniques seen in extremist recruitment, and the analogue intervention *Fake News* aims to inoculate against the misinformation techniques used in misinformation on immigration.

Game Genre

I have identified three broad categories of game genre in this review (linear roleplaying text adventure games, quiz games, and interventions with unique designs), overlapping with previous wider reviews on misinformation games more generally (DeJong, 2023; Kiili et al., 2024). *Bad News*, *Harmony Square*, *Go Viral!*, *ChamberBreaker*, *Misinformation Is Contagious*, *COVID-19 Fake News*, *Join This Group*, *Radicalise*, and *Bad Vaxx* (both editions) have all been designed with largely similar gameplay mechanics, and make this game genre the most popular in game-based misinformation interventions at time of writing. These games are perspective-taking exercises that include a simple storyline or

progression, where the player often manages a resource (e.g., credibility, followers, or reliability) while disseminating information (posting, direct messaging or publishing). These games are fundamentally text adventure games: despite some visual elements, players interact with the game primarily through a text-based narrative. Some games give a sense of jeopardy during play through positive/negative feedback on player choices, but none of the games reviewed in this category will allow the player to 'lose', following a pre-defined linear storyline.

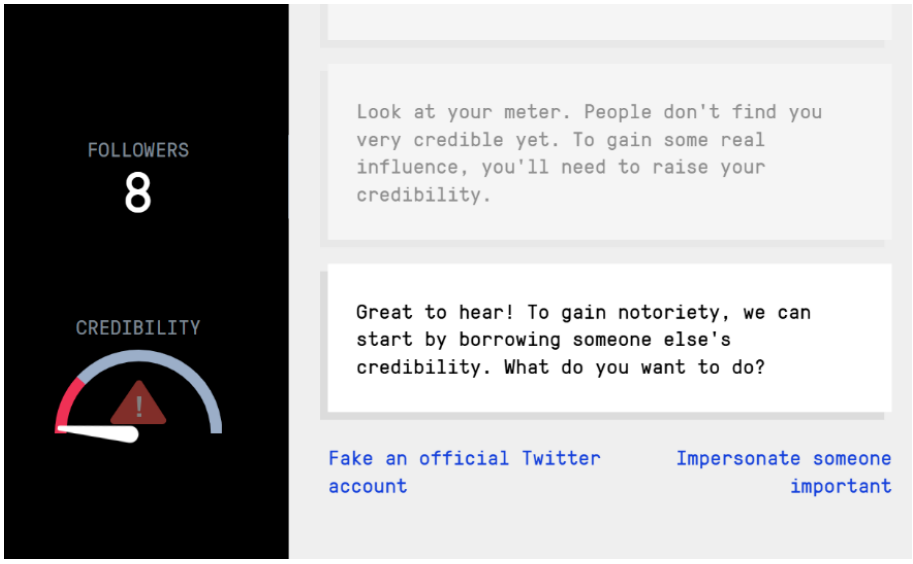


Figure 22: The Bad News game (Roozenbeek & van der Linden, 2019): an early digital game-based inoculation intervention to combat misinformation.

The second largest category of games reviewed are quiz games. Fundamental quiz-based interventions such as the *Spot The Troll* and *COVID-19 Trivia* have players identifying headlines and content as being either real/true or fake/misinformation. Players are often educated of background information and misinformation techniques used in specific quiz questions and are given a final score on completion. Some game-based interventions use quiz elements alongside other game genres. For example, *Cranky Uncle* and *Cranky Uncle Vaccine* interlace quizzes with narrative and informational segments throughout, introducing players to tested concepts in advance.

The remaining games not yet discussed are categorically unique. The game *Fake News* is a collaborative phrasal template card game where players work together to write an article using template cards to match a given character. *Fake News Detective* – a role-playing adventure game – combines heavy narrative elements and quizzes, with the player taking the perspective of a detective trying to find a 'misinformation boss'. The third of three microgames by Grace et al. (2024) focusses heavily on narrative elements, with the player supporting a friend who "may be making unhealthy, mis/disinformed decisions regarding their immunocompromised health" (Grace et al., 2024, p. 5517). *Cat Park* is also heavily focussed on narrative, with the player sowing

dissent towards the building of a cat park through misleading headlines, memes, and altering images.

Game Platform

The platform of a game describes the software, device, or setting through which a game can be played. From the games reviewed in this article, there is a heavy partiality towards digital interventions: 18 of the 25 interventions are digital, 15 of which are designed only for browser. The remaining three digital interventions (all three editions of *Cranky Uncle*) can be played through both the browser and mobile. Of the four analogue interventions, *Fake News* and *Follow Me* are card-based games, and *Werewolf On Campus* is a game-based classroom activity.

Target Audience

Reviewing target audiences of games-based interventions, I consider if the intervention has been tailored for a single, specific audience. Of the 25 games covered in this review, seven have explicitly indicated as having made design decisions to focus on a particular demographic. These games are targeted for children (*Unnamed 3*, *Follow Me*, F*ake News Detective*), players in countries such as Iraq (*Radicalise, Iraq edition*), North India (*Join This Group*), and countries in East and West Africa (*Cranky Uncle Vaccine*). Table 1 defines the target audiences for the game-based interventions in this review; however, I do not infer target audience from study sample demographics as these may in some cases be convenience samples and not indicative of the game designer's intentions.

Roleplaying Perspective

Despite being a relatively small and specific part of game genre, the perspective the player takes in roleplaying games (i.e., playing from an 'bad' perspective spreading misinformation, or a 'good' perspective defending against misinformation) has been the subject of recent research and discussion (Appel et al., 2025; Harjani et al., 2023). Of the game-based interventions reviewed, nine give the player a 'bad' perspective disseminating misinformation, and eight give the player a 'good' perspective fighting misinformation. *Cat Park* is the only game-based intervention I have identified to integrate both perspectives into the game, with the player's perspective changing during the game's narrative.

The interventions reviewed demonstrate a promising coverage of topics and real-world problems by scholars. Topics that often target by misinformation, commonly health and elections (Stepney & Lally, 2024) have been well represented in the reviewed game-based interventions. Game-based interventions have also sought to inoculate against different types of misinformation, including misinformation that is multimodal (*CatPark*), message-based (*Join This Group*), and echo chambers (*ChamberBreaker*). Other disciplines have in some cases also drawn from game-based

inoculation literature, creating interventions on topics tangential to misinformation (see *Spot The Troll*, *Radicalise*, and *Unnamed 3*). This demonstrates the accessibility of inoculation as a communication theory to underpin intervention design. Nonetheless, conflict and artificial intelligence (reoccurring and emerging topics in misinformation research, see Stepney & Lally, 2024; Alkaissi & McFarlane, 2023; Gravel et al., 2023) have received little attention from game-based inoculation scholars. How game-based approaches perform in environments where disinformation is an offensive combat tactic (Doroshenko & Lukito, 2021; McGeehan, 2018; Roozenbeek, 2024) is also under-explored (but see Saleh et al., 2023).

However, this section finds little exploration of platforms and how platforms may differently moderate effectiveness. The reach and scalability of online interventions have allowed faster large-scale global testing particularly with quantitative research (Grace & Hone, 2019; Maertens et al., 2025; Roozenbeek & van der Linden, 2019). Despite this, analogue interventions offer a very different experience for players, with the potential to stimulate increased enjoyment, retention, and interest in players (Adams, 2023; Axelsson et al., 2024). Some serious game designers have also suggested that the social characteristics of analogue games, for instance, may better stimulate learning outcomes in some environments (Graffer et al., 2015; N. Henderson et al., 2024). Yet, only four interventions assessed in this review are analogue in design. Furthermore, these interventions have shown mixed results (Pomichal & Trnka, 2023; Rędzio et al., 2023; Roozenbeek & van der Linden, 2018), suggesting that further investigation and qualitative observation is required to understand nuances in analogue intervention design and its impact on intervention effectiveness. Furthermore, there is little research understanding the potential for platforms such as virtual reality and escape rooms to host interventions against misinformation (but see Erisen et al., 2024; Paraschivoiu et al., 2021).

This section also finds little exploration of game genre. One of the first game-based interventions to combat misinformation using inoculation theory was *Bad News*: a menu-driven narrative-based text adventure game (see Figure 22). The success of the game *Bad News* has likely inspired further game-based interventions; I find that 10 further interventions in this review share similar game genre design (an observation also made in other reviews, see Kiili et al., 2024). This may be a symptom of a lack of understanding in how to design effective game-based interventions for specific topics and audiences, leading practitioners to adapt previously successful interventions to new applications instead of making new empirically backed design decisions. Although it is not necessarily the case that different individual game mechanics are better for achieving intervention effectiveness (game design literature stresses that game mechanics are not solitary but work together, being highly sensitive to change in one

another, see Fullerton, 2004), it may be that some broader genres or platforms (such as multiplayer design) can help to build resistance differently.

This section has reviewed the design trends of 25 game-based misinformation inoculation interventions, including variables on topic, genre, platform, target audience, and roleplaying perspective. The section finds that, although interventions cover a wide range of topics and collaboration from related disciplines, there is an under-exploration of many design aspects. Most interventions reviewed are digital, with very little research exploring the affordances of analogue or virtual reality platforms. Furthermore, this review has revealed that many digital interventions share similar design to *Bad News*, despite there often being little empirical justification for adopting this design for the new context.

6.2 Interpretation of Inoculation Theory in Game-based Misinformation Interventions

Despite the recent popularity of game-based approaches in inoculation research, no single review has categorised how these interventions elicit inoculation theory mechanisms. A review is an essential epistemic tool in science research that can not only provide a summary of existing research but can help guide scholars with future work (Palmatier et al., 2018). Discerning how game-based interventions draw on inoculation theory can not only establish consensus of approach but can highlight areas of inoculation literature that have been underexplored in this context. This section reviews 25 game-based inoculation interventions, summarising different approaches to eliciting inoculation mechanisms such as threat and defensive counterarguing, and contemporary ideas in the inoculation space such as feedback exercises. Table 9 summarises the findings related to inoculation theory variables presented in this section.

*Table 9: Analysis of inoculation theory intervention design variables in game-based interventions * indicates that the entry is an assessment made by the author of this review. N.D. denotes where a variable is not discussed in the corresponding paper, and the intervention is not publicly available for playtesting.*

Intervention name	Paper Id	Threat source	Counterarguments type	Explicit forewarning component	Debriefing component
Fake News	1	Implicit	Implied technique-based	No	No
Bad News	2	Implicit*	Technique-based	No	Throughout (informational) and at end (informational and quiz task)
Harmony Square	3	Implicit*	Technique-based	No	Throughout (informational)

					and at end (quiz task)
Chamber-Breaker	4	Implicit*	S.D. (assessed technique-based)	No	N.D.
Go Viral!	5	Implicit	Technique-based	No	Throughout (informational) and at end (informational and quiz task)
Cranky Uncle	6	Explicit forewarning & implicit*	Logic-based	Yes	Throughout (quiz tasks)
Join This Group	7	Implicit*	Technique-based	No	Throughout (informational)
Fake News Detective	8	Explicit forewarning	N.D. (assessed technique-based)	Yes	Throughout (quiz tasks) and at end (informational)
Cranky Uncle Vaccine (East Africa edition)	9	Explicit forewarning & implicit*	Logic-based	Yes	Throughout (quiz tasks)
Misinformation Is Contagious	10	Implicit*	Technique-based	No	N.D.
Follow Me	11	Implicit*	N.D. (assessed technique-based)	No	Interactive teacher-led debrief encouraged
Spot The Troll	12	Implicit*	S.D. (assessed technique-based)	No	Throughout (informational) and at end (informational)
COVID-19 Fake News (CFN)	13	Implicit*	N.D. (assessed technique-based)	No	N.D.
Unnamed 1	14	Implicit*	N.D. (assessed technique-based)	No	Throughout (informational)
Radicalise (Iraq edition)	15	Implicit*	Technique-based	No	N.D.
Werewolf On Campus	16	Implicit*	N.D. (assessed technique-based)	No	Informational debrief at end by researcher
Cat Park	17	Implicit*	Technique-based	No	Reviewed through narrative
Covid-19 Trivia	18	Implicit*	N.D. (assessed technique-based)	No	None
Myth busters	18	Implicit*	N.D.	No	N.D.
Unnamed 2	18	N.D.	N.D.	No	N.D.

Radicalise	19	Implicit*	Implied technique-based	No	N.D.
Cranky Uncle Vaccine (Ghana edition	20	Explicit forewarning & implicit*	Logic-based	Yes	N.D.
Unnamed 3	21	N.D.	N.A.	No	None
Bad Vaxx (good edition)	22	Explicit forewarning	Technique-based	Yes	Throughout (informational)
Bad Vaxx (evil edition)	22	Explicit forewarning	Technique-based	Yes	Throughout (informational)

Eliciting Inoculation Theory Mechanisms: Threat

Threat is a central mechanism to inoculation theory and differentiates the theory from other approaches such as two-sided messaging and simple one-sided pre-bunking (Compton, 2024). When reviewing game-based inoculation interventions against misinformation, I have considered not only the design of the interventions, but also how these interventions stimulate this threat mechanism in the player. In conventional inoculation messaging threat is elicited through either/both a forewarning and implicitly through a refutational pre-emptive task. This is achieved similarly in the reviewed game-based interventions, being elicited (a) through an explicit in-game forewarning, and/or (b) implicitly through game content. Some papers explicitly state this: papers 22, 20, and 8 (*Bad Vaxx*, *Cranky Uncle Vaccine*, and *Fake News Detective*) are described as delivering threat primarily through an explicit forewarning during the prologue of the game; *Fake News* and *Go Viral!* are described as delivering threat implicitly through the content of the intervention itself (see implicit threat, Compton, 2021). Overall, only six (out of 25) game-based interventions have corresponding papers that explicitly state how threat is elicited in players.

Findings from the playtesting of these game-based interventions has not always concurred with the assessment made by authors in these corresponding papers. For example, the intervention *Fake News Detective* is said to begin with a forewarning message "telling the player that the world is facing a serious threat from fake news" (Tandoc & Seet, 2023, p. 774). This forewarning (and other interventions claiming to have a forewarning, see Table 10) does little to alert the player of their attitudinal or behavioural vulnerability, and is in all likelihood something they already know (Altay, 2023). A more effective forewarning message strategy may, for example, warn players that being complacent when reading new information may lead to the belief of misinformation, or explain the cognitive and motivational biases that prevent individuals from being able to unbiasedly evaluate new information (for inoculation component design see Ivanov, 2012, 2017; Parker et al., 2022).

Table 10: Inoculation forewarning messages at the beginning of three game-based interventions reviewed. Forewarnings are split into individual bullet points in the order in which they are displayed to the player.

Fake News Detective	***Bad Vaxx (both versions)***	***Cranky Uncle Vaccine***
• The world is facing a new terror: fake news. • Police are looking for the mastermind behind the spread of fake news in the online world. • You have been recruited to be a spy to find and catch the mastermind. • You have been wearing a disguise to hide your identity during the investigation. • Finally, after weeks of investigation and following various leads, you found the headquarters where fake news is being produced. • This is your chance to finally catch the mastermind.	• In this game you're gonna use misinformation tactics yourself. This way, you'll get better at spotting them in the wild, kind of like a psychological 'vaccine' against manipulation.	• Hi! I'm a nurse. I'm here to help you understand how vaccines save lives. • And I'm a science denying Cranky Uncle who doesn't believe in vaccines. • There's lots of science to show vaccines are safe. It's a fact! • I don't care about facts. I'll show you the different tricks I use to cast doubt on facts. • As you learn new tricks, you'll earn cranky points and see me get crankier! • When you understand Cranky Uncle's tricks, you're less likely to be fooled.

Even where not explicitly stated in a corresponding paper, it can be assumed that all 25 game-based interventions deliver threat implicitly to a certain degree. A forewarning is not a necessary component in conventional and game-based interventions (Compton, 2024; Ivanov, 2017); in fact, early work on inoculation did not include forewarning components at all (McGuire, 1961). One design aspect discussed as eliciting threat in game-based interventions is the roleplaying perspective taken by the player (Appel et al., 2025). It is hypothesised that taking on a 'bad' perspective may elicit more threat than a 'good' perspective "because it may make people feel slightly threatened and uncomfortable about their in-game actions" (Appel et al., 2025, p. 2). Unfortunately, no evidence yet exists comparing the two approaches in relation to threat. Paper 1 discusses how its analogue intervention the *Fake News* game elicits threat through negative affective content, drawing on research by Pfau et al. (2009). Despite this, exactly how the reviewed game-based interventions deliver threat to the player, and the degree to which this is successfully achieved, has seen very little attention. Moreover, the type of threat delivered through game-based interventions (i.e., apprehensive or motivational) has seen little attention. Paper 14 investigated the threat mechanism in game-based interventions found that the intervention did not increase measured threat in participants (Ma et al., 2023).

Eliciting Inoculation Theory Mechanisms: Defensive Counterarguing

To review, defensive counterarguing (the second core mechanism in inoculation theory) is motivated and/or improved through different components depending on if

the intervention is issue-based or technique-based (see Section 2.2). Issue-based inoculation improves defensive counterarguing by providing individuals with counter-attitudinal arguments paired with corresponding refutations. This provides the individual with counterarguments that can be used when subsequently under attitudinal attack. None of the game-based interventions reviewed claim to be exclusively issue-based. However, after playtesting, this review assesses that *Covid-19 Trivia* is in fact issue-based. As a quiz-style game, *Covid-19 Trivia* presents arguments on COVID-19 and vaccines, as well as communicating the veracity of these arguments.

Rather than raising single counter-attitudinal arguments and refutations, technique-based inoculation informs individuals of the rhetorical and malicious techniques used in the generation of misinformation content, with the aim of conferring resistance to a wider range of content than a single issue (Cook et al., 2017; Roozenbeek & van der Linden, 2018, 2019). This review finds that technique-based inoculation represents the fundamental actualisation of inoculation theory for many of the games in the review; *Fake News*, *Bad News*, *Harmony Square*, *Go Viral!*, *Misinformation is Contagious*, *Join This Group*, *Radicalise (both editions)*, and *Cat Park* are described in their corresponding papers as being technique-based. *Cranky Uncle* and *Cranky Uncle Vaccine (both editions)* are described as logic-based interventions, a concept closely related to technique-based inoculation.

The technique- and logic-based interventions in this review have in many cases used existing misinformation frameworks as the basis for their intervention content. *Fake News*, *Bad News*, *Harmony Square*, *Cat Park*, *Go Viral!*, *Misinformation Is Contagious*, and *Cranky Uncle Vaccine* inoculate with frameworks by Agarwal et al. (2017), Barzilai and Chinn (2018), van der Linden and Roozenbeek (2020), and World Health Organisation (2020). Other games such as *ChamberBreaker*, *Spot The Troll*, *Join This Group*, *Radicalise*, and *Bad Vaxx* use techniques compiled from separate sources. With the exception of *Cranky Uncle* and *Cranky Uncle Vaccine (both editions)*, and *Harmony Square* and *Cat Park*, no two of the gamified interventions reviewed share a single framework of misinformation techniques.

Of the 25 game-based interventions reviewed, *Bad Vaxx* is the only to claim to combine both approaches: "The theory-guided design of the Bad Vaxx game combines the strengths of issue-based and technique-based inoculation […] while the game focuses on the issue of vaccine misinformation, it successfully teaches participants techniques that are commonly used to spread different kinds of misinformation rather than specific counterarguments" (Appel et al., 2025, p. 10). This conceptualisation of issue-based inoculation is contrary to conventional inoculation literature where pre-emptive refutational tasks (i.e., interacting with counter-attitudinal arguments and respective refutations) are core markers of issue-based inoculation (not just inoculating

on a topic, see Compton, 2012, 2024; Roozenbeek, Traberg, et al., 2022). Given the lack of clarity in this area of the literature, the author believes that, although *Bad Vaxx* inoculates individuals on a single issue (i.e., vaccine misinformation), it should be defined as a technique-based intervention.

Contemporary Inoculation Message Components: Feedback Exercises

Although not part of McGuire's original inception of inoculation theory, scholars have recently investigated how the integration of new message features can complement and bolster (game-based) intervention effectiveness. Recent research in inoculation theory has found that feedback exercises (i.e., providing information regarding the individual's performance or understanding) can have a significant positive effect on individuals, and could help reduce decay of conferred resistance (Hattie & Timperley, 2007; Leder et al., 2024). Interventions reviewed in this chapter include either (a) no debriefing or feedback component in the intervention, (b) a summary of key concepts inoculated on, or (c) a quiz for players to test their understanding, giving feedback on answers. Of the 25 game-based interventions reviewed, six integrated additional (excluding quiz games themselves) quiz-style feedback exercises. Six games also integrated textual debrief material, summarising content that players had learned in the intervention. *Follow Me* and *Werewolf On Campus* are designed to support educator-run feedback sessions. This review does not consider methodological post-test evaluation tasks as being feedback exercises as they are not inherently part of the intervention's design (but see Maertens et al., 2021).

In the previous section of this chapter I found that game-based interventions against misinformation cover a promising range of topics and real-world issues. In this section I similarly find a diverse range of misinformation frameworks from which technique-based interventions inoculate on. This is promising, as serious game design literature advises caution in serious game developers using their own knowledge rather than drawing from literature to inform the content of their game (Culyba, 2018). 17 out of the 25 game-based interventions drew on exiting research on manipulative or rhetorical misinformation techniques. How research from these frameworks is communicated through game-based design has seen little interrogation, however, and future research should seek to understand when specific frameworks are more effectively communicated (e.g., different levels of abstraction in misinformation technique frameworks may be more appropriate given a target audience's prior knowledge, see Kalyuga et al., 2003).

However, from the 25 reviewed game-based interventions citing inoculation as their foundational theory, I find a distinct lack of clarity and discussion of how inoculation's core mechanism – threat – is elicited. Only six interventions explicitly describe how threat is elicited from players, despite threat being a required feature in

inoculation interventions (Compton, 2024). That the mechanism through which threat is elicited in three game-based interventions has only been identified in subsequent publications (Appel et al., 2024; Cook et al., 2024), by (in some cases) different authors, further highlights a lack of clarity in the literature. Indeed, how game-based interventions are claimed to elicit threat also differs between authors, being (a) through an explicit forewarning at the beginning of the game, (b) implicitly through game content, such as "negative affective content" (Roozenbeek & van der Linden, 2018, p. 572) or roleplaying narrative game design.

In this review section I have assessed 25 game-based inoculation interventions building players' resistance against misinformation. I find diversity in intervention topics and frameworks used in technique-based interventions, but homogeneity in platform and genre. I also find that information on design processes and consideration of inoculation theory was often either superficially described or not described at all in corresponding papers. Many of the assessments made in this review are therefore subjective judgements from the author (informed through playtesting), but this has been particularly difficult where the intervention itself is not publicly available. I have reviewed these interventions against eight variables; however, I encourage readers to review the full dataset included in the supplementary materials at https://doi.org/10.6084/m9.figshare.29310692, which include analysis of a further five variables. In brief, this review finds that discussion and justification of intervention design decisions, and exactly how inoculation theory had been considered in design, is lacking.

6.3 Is Inoculation Still an Appropriate Label for Reviewed Game-based Interventions?

In this review I find that discussion of how inoculation theory has been integrated into the design of game-based interventions—other than the fact that the interventions are designed to pre-emptively delivered—is often superficial or lacking. Very rarely is inoculation theory demonstrated as being a driving concept informing the design of the intervention. This may be a contributor to why there is relatively little research understanding underlying inoculation theory mechanisms in game-based contexts. In this section I review the findings of the previous section in relation to the inoculation theory literature as a whole and discuss the underexplored link between inoculation theory and the wider serious game design literature. Table 11 summarises these findings.

*Table 11: Described threat source (i.e., implicit or explicit), described design considerations taken to elicit this threat, and intervention efficacy in eliciting this threat. * indicates where the corresponding paper does not present any quantitative results on intervention efficacy. N.D. denotes where a variable is not discussed in the corresponding paper.*

Paper Id	Intervention name	Described threat source	Described design considerations specific to inoculation theory	Measured increase in threat
1	*Fake News*	Implicit	Negative affective content designed to increase threat	Not measured
2	*Bad News*	N.D.	N.D.[16]	Not measured
3	*Harmony Square*	N.D.	N.D.[16]	Not measured
4	*Chamber-Breaker*	N.D.	N.D.	Not measured
5	*Go Viral!*	Implicit	Narrative designed to increase motivational threat[16]	No significant change in apprehensive threat observed, but marginally significant motivational threat observed
6	*Cranky Uncle*	N.D.	N.D.	Not measured*
7	*Join This Group*	N.D.	N.D.	Not measured
8	*Fake News Detective*	Explicit forewarning	N.D.	Significant increase in threat observed
9	*Cranky Uncle Vaccine (East Africa edition)*	N.D.	N.D.	Not measured
10	*Misinformation Is Contagious*	N.D.	N.D.	Not measured
11	*Follow Me*	N.D.	N.D.	Not measured*
12	*Spot The Troll*	N.D.	N.D.	Not measured
13	*COVID-19 Fake News (CFN)*	N.D.	N.D.	Not measured
14	*Unnamed 1*	N.D.	N.D.	No significant change in apprehensive threat observed
15	*Radicalise (Iraq edition)*	N.D.	N.D.	Not measured
16	*Werewolf On Campus*	N.D.	N.D.	Not measured

[16] The interventions *Bad News*, *Harmony Square*, and *Go Viral!* have subsequently been described as eliciting motivational threat through a 'bad' roleplaying perspective by Harjani et al. (2023).

17	*Cat Park*	N.D.	N.D.	Significant increase in motivational threat observed
18	*Covid-19 Trivia*	N.D.	N.D.	Not measured*
18	*Myth busters*	N.D.	N.D.	Not measured*
18	*Unnamed 2*	N.D.	N.D.	Not measured*
19	*Radicalise*	N.D.	N.D.	Not measured
20	*Cranky Uncle Vaccine (Ghana edition)*	N.D.	N.D.	Not measured
21	*Unnamed 3*	N.D.	N.D.	Not measured
22	*Bad Vaxx (good edition)*	Explicit forewarning	N.A.	Not measured
22	*Bad Vaxx (evil edition)*	Explicit forewarning	'bad' perspective taking to increase threat	Not measured

The application of game design and inoculation theory to fight misinformation represents an interdisciplinary collaboration between the communication theory and serious game literature, yet many of the corresponding papers of reviewed interventions do not effectively draw on inoculation theory. Using inoculation theory in game-based contexts has produced effective interventions, (Basol et al., 2021; Cook et al., 2023; Lees et al., 2023), but there is much still to be learned. The nature of game-based misinformation inoculation interventions is very different to the pioneering vision of McGuire's original reading and writing tasks and may stimulate very different cognitive and affective processes. One should not, for instance, assume that research and findings understanding the mechanisms of inoculation theory will generalise to these game-based contexts. In this review, I find that the actualisation of inoculation theory in these game-based context has gone broadly underdiscussed in the reviewed literature. For example, only 4 papers discuss the type of threat the intervention is designed to elicit, and only 3 papers describe design considerations taken to achieve this. This review finds that many authors of the interventions' corresponding papers draw from the inoculation theory literature superficially, engaging primarily with the inoculation metaphor and not the underlying mechanisms of the theory itself.

Cranky Uncle, *Cranky Uncle Vaccine (both editions)*, *Bad Vaxx (both editions)*, and *Fake News Detective* are all described in writing as delivering threat through an explicit forewarning during the prologue of the game. However, Section 6.2 discusses how these forewarning messages may be largely ineffective at eliciting any new sense of threat in the player and do not follow established forewarning design protocol (Ivanov, 2017). Furthermore, these forewarning components appear to be no more substantial than in other reviewed interventions such as *Join This Group* and *Harmony Square* (which are not

described as having a forewarning component). As highlighted in Section 6.2, one cannot assume that forewarning components remain effective approaches to eliciting threat in game-based contexts. Game length of the reviewed interventions varied from 15 minutes (*Join This Group*) to 45 minutes (*Cranky Uncle Vaccine*) or more (*Fake News Detective*). It is unclear how impactful an explicit forewarning can be if it is immediately followed by a 30-minute serious game (as opposed to a pre-emptive refutational task, as with a conventional issue-based inoculation).

This chapter has also reviewed how threat can be delivered *implicitly* through game content. Conventional inoculation literature indicates that, as well as arming an individual with refutations to counterarguments they are likely to encounter, refutational pre-emption can elicit its own implicit threat (although this is not always to the same strength without an explicit forewarning, see Compton & Ivanov, 2012). It is intuitive therefore that the reviewed games should also elicit some implicit threat from the player, and authors have suggested that this is achieved through negative affective content or a 'bad' roleplaying perspective. However, only papers 5, 8, and 14 measure changes in threat (either apprehensive or motivational) signalling a lack of evidence on the moderating effect of this variable in game-based contexts. Findings from these papers are mixed, and there is some evidence to suggest that taking a 'good' roleplaying perspective (note that 'bad' roleplaying perspective taking has been discussed as an approach to increasing threat) may be more effective (Appel et al., 2025).

The potential for game design elements such as 'bad' roleplaying perspectives to elicit threat is further complicated by competing communication theories such as transportation theory. Transportation theory posits that an effective narrative can change the attitudes and beliefs of the person consuming that media (M. C. Green et al., 2004). Effective narrative can elicit empathetic sympathy from the player, and can transport the player "to the values, likes, and dislikes of the story's portrayals" (Grace & Liang, 2023, p. 4693). Application of transportation theory to this game-based context has seen players playing as, or empathising with, individuals effected by misinformation (see Grace & Liang, 2023). There is clear conflict between this communication theory and 'bad' perspectives taken in the games discussed above, and thus one must question existing literature on which approach is more effective. Other design elements used in many of the reviewed interventions are also contrary to inoculation theory mechanisms. Some scholars have integrated elements of humour into game-based interventions aiming to make "the issue less threatening and more accessible" (Cook, 2021, p. 14). Although this may reduce reactance from players with opposing initial attitudes or beliefs (a common justification for this approach, see Roozenbeek & van der Linden, 2020), this may in fact reduce the apprehensive threat elicited implicitly through a 'bad' roleplaying perspective (or otherwise).

Inoculation theory scholars have more recently suggested that threat should be considered to be more motivational (rather than apprehensive) in nature, as the threat mechanism itself is posited to motivate defensive counterarguing (Banas & Richards, 2017; Maertens et al., 2025). This conception of threat is more congruous with the design of current game-based interventions (i.e., intervention content is broadly designed to improve defensive counterarguing against misinformation, for discussion see Roozenbeek & van der Linden, 2018; Saleh et al., 2023), but discussion on this is very limited. Research has measured motivational threat through participants self-reporting motivation to defend an attitude (Basol et al., 2021; Neylan et al., 2023), but this may not translate to real world counterarguing (Neylan et al., 2023). Nonetheless, both papers that measure motivational threat (see papers 5 and 17) have shown promising results.

A clear characteristic of the games and corresponding research covered in this review is the absence of any formalised approach to using inoculation as a foundational theory for game-based interventions. For conventional, issue-based inoculation the structure is clear: a forewarning component helps to deliver threat, and refutational pre-emption raises and refutes inaccuracies on the issue (as well as implicitly delivering threat). For game- or technique-based inoculation, threat can be delivered explicitly or implicitly through a forewarning, perspective taking, or game content, and refutational pre-emption takes the form of communication of the techniques chosen to inoculate the player on. This game-based design of intervention, then, shares very few similarities to the conventional two-component inoculation message. It is not unexpected, therefore, that scholars have expressed frustration at not knowing "what works" in game-based intervention design (Appel et al., 2025; Harjani et al., 2023; Ma et al., 2023; Roozenbeek et al., 2024; Saleh et al., 2023), particularly when null results are interpreted and analysed in relation to conventional inoculation mechanisms (Harjani et al., 2023; Ma et al., 2023; Neylan et al., 2023).

One must ask, therefore, if the theoretical reliance on inoculation theory observed in the reviewed research is still warranted, as in reality many game-based interventions are struggling to effectively draw from the theory at all. Inoculation literature has helped practitioners and scholars create effective conventional interventions (Banas & Rains, 2010) by informing their design (Compton, 2012; M. L. M. Wood, 2007). Conversely, inoculation literature offers very little instructive support for the creation of game-based interventions, and conventional inoculation theory components and mechanisms have received little attention in this context. Other wider reviews of misinformation serious games (see DeJong, 2023; Kiili et al., 2024) identify many game-based interventions that do not draw on inoculation theory; yet, these share many similarities to the interventions reviewed in this chapter. For example, the

serious game *Fakey* (Micallef et al., 2021) tasks players with reviewing the veracity posts on a fictitious social media platform. A key difference between this serious game and those reviewed is that the reviewed interventions are said to be grounded in inoculation theory. Yet, *Fakey* shares many close similarities to minigame tasks in *Fake News Detective*, one of the game-based interventions reviewed in this chapter. The findings of this chapter suggest a pressing need for scholars creating game-based misinformation interventions to more clearly identify and operationalise the underlying theory, which may or may not be inoculation theory.

This section has analysed and discussed the findings from this review relating to inoculation theory. Specifically, the section has interrogated the interpretation of inoculation theory in game-based contexts, arguing that not only are many of the reviewed interventions ineffective at eliciting threat (apprehensive or motivational), but inoculation mechanisms have gone underexplored and underdiscussed in game-based contexts. This section concludes by considering inoculation theory's role in the reviewed interventions.

6.4 Identifying Underexplored Areas of Game Design and Discussion

Within the game-based misinformation interventions assessed in this review and their corresponding papers, I found not only a lack of engagement with key inoculation theory mechanisms and components, but a lack of exploration in intervention design (i.e., in game genre and platform). In the following section I review the design processes undertaken by the 25 game-based interventions, finding little discussion of design processes and justification for design decisions. I position these findings within the context of the previous two chapters and discuss how inoculation theory could be integrated with serious game design literature to provide practitioners with a more instructive design processes than inoculation alone.

The design process of a (serious) game refers to the methodological steps taken to create the game itself (Fullerton, 2004; Schell, 2019). A notable finding from this review is a lack of discussion surrounding design processes in the games' corresponding papers. Given that some scholars have expressed uncertainty surrounding "what works" in game-based intervention design (Appel et al., 2025; Basol et al., 2021; Saleh et al., 2023) it is surprising that this has received little attention. Some recent research has sought to determine whether design decisions relating to the roleplaying perspective taken by the player affects the interventions efficacy (Appel et al., 2025). Game design literature, however, stresses that designed mechanics are not solitary, but work together, being highly sensitive to change in one another (Fullerton, 2004; Salen & Zimmerman, 2004; Schell, 2019). Although focussing on individual components in

conventional inoculation interventions has been productive, game-based interventions are significantly more complex, and the game design philosophy surrounding design processes may be particularly relevant here. Instead of focussing on individual design elements of interventions, scholars should interrogate the design *processes* taken to achieve the most effective serious game. Discussion of design processes adopted by practitioners designing game-based interventions may be an important step to creating more effective interventions.

This review also finds little discussion of design considerations taken to tailor interventions to their target audiences or intended application. This lack of enquiry may be partly responsible for null results in some interventions reviewed here. For example, the game *Join This Group* was designed by using similar mechanics to *Bad News* but was created for a different subtopic of misinformation (direct messaging), and for a different audience (adults in broadly rural Northern India). Given *Bad News's* promising results, it was surprising that *Join This Group* had not been successful (Harjani et al., 2023). Harjani et al. (2023) themselves reflect that a greater consideration of the cultural differences of the target audience may have resulted in a more effective intervention.

The findings from this review indicate that the findings with the intervention *Fake News* in the previous chapter (i.e., that many attitudinal changes observed in players could be equally explained by learning theories than inoculation theory) may extrapolate to other game-based interventions as well. Chapter 5 argues that game design decisions of the game-based intervention *Fake News* have an important impact on intervention outcomes. Chapter 5 also finds that the intervention *Fake News* was ineffective at eliciting threat due to players' prior knowledge, and that a more considered integration of inoculation theory was required. The review's findings on inoculation theory components and mechanisms supports this as well. Findings in this review suggest that the calls heralded in the previous chapter may therefore be applicable to the wider game-based inoculation field: more focus on learning, better consideration of how inoculation theory mechanisms are elicited, and more interaction with (serious) game design literature are required.

Game design scholars and practitioners have long established effective formal processes to creating enjoyable games, including instructional design methods and frameworks that help guide practitioners in the game design processes (Ávila-Pesántez et al., 2017; Laine & Lindberg, 2020; Mora et al., 2015). Central to many of these approaches is the philosophy of iterative design. Instead of designers working towards a single finished game, it is widely accepted that it is more effective for designers to create prototypes, playtest these prototypes (i.e., test the efficacy or success of the game with a small sample of players), and revise based on results/feedback throughout development (Culyba, 2018; Fullerton, 2004; Salen & Zimmerman, 2004; Schell, 2019).

This 'playcentric' approach (for further discussion on playcentric design, see Fullerton, 2004) encourages repeated interrogation of game design decisions throughout the development process, allowing designers to curate and adjust a game's design until the game reaches a satisfactory level of quality (or the project is limited by other constraints such as time or budget). This iterative design process is not exclusive and underpins many other (serious) game design frameworks (e.g., Alvarez et al., 2019; Culyba, 2018; Flanagan et al., 2005; Norman, 2024).

This is not to say that the games reviewed in this chapter have not been designed following formal game design processes. Many of the game-based interventions may have been designed and developed not directly by authors of reviewed papers (although authors may have been involved in the design process), but by established game development studios (such as Gusmanson and Goodbeast, see Table 7). However, the lack of transparency surrounding design processes restricts practitioners from building upon previous research and may result in them adopting ineffective design processes. Clarity and transparency of design processes may enable scholars to adopt, adapt, and build upon existing approaches, which can ultimately result in more effective game-based interventions against misinformation.

Promising research by Cook et al. (2024) and Hopkins et al. (2023) has demonstrated the value of discussion and enquiry into game design processes, particularly in relation to target audiences. Through co-design workshops with in-country partners and stakeholders, Cook et al. (2024) and Hopkins et al. (2023) directly informed the content and design of the game *Cranky Uncle Vaccine*, altering text translations, character clothing, and character skin tone to be more regionally interpretable and accessible. For example, Figure 23 shows the different character designs for the deployment of the *Cranky Uncle Vaccine* game in Ghana, Uganda, Kenya, and Rwanda. Empirical testing with the *Cranky Uncle Vaccine* game showed a significant reduction in vaccine hesitancy from players, with individuals more likely to get vaccinated after playing the game (Cook et al., 2024).

Figure 23: Designs of the 'Cranky Uncle' character. Left: 'Cranky Uncle' character from the Cranky Uncle game (from which Cranky Uncle Vaccine is designed, see Cook et al., 2023). Centre: health worker/nurse and 'Cranky Uncle' from the Cranky Uncle Vaccine game, designed for Ghana (see Cook et al., 2024). Right:

health worker/nurse and 'Cranky Uncle' from the Cranky Uncle Vaccine game, designed for Uganda, Kenya, and Rwanda (see Hopkins et al., 2023).

More transparency surrounding design rationale and processes may also help practitioners successfully adopt a more theory-driven design approach. The previous section of this chapter reviewed how, although scholars intended for inoculation to be a foundational theory for their serious games, many of the reviewed games struggled to effectively elicit core inoculation theory mechanisms. Moreover, design decisions (e.g., humorous design, and 'good' perspective taking) often conflicted with the designer's conceptualisation of the threat mechanism (e.g., threat as apprehension). Given these findings, one must question whether the effectiveness of the reviewed game-based interventions is due to inoculation being their foundational theory, or instead due to successful serious game design. Importantly, this review does not argue that inoculation theory is ineffective in game-based contexts. Instead, I argue that scholars and practitioners should more closely study how inoculation theory can be considered in the design processes and discuss this with greater transparency.

This section assesses how design rationales and processes are presented in the reviewed game-based interventions, finding that of the reviewed games' corresponding papers the design processes and design considerations are underdiscussed or in some cases not discussed at all. Given that the previous section finds that many reviewed interventions do not effectively draw from inoculation theory, this section argues that effective (serious) game design processes may in fact be responsible for the success of many game-based interventions reviewed. (Serious) game design literature places great importance on particularly iterative design processes to create effective and successful (serious) games. For these reasons, this section argues that more transparency and deeper interrogation of design rationale and processes are essential to create more effective game-based interventions against misinformation.

Summary and Conclusion

Reviews of misinformation serious games have been completed before (see DeJong, 2023; Kiili et al., 2024), but these have focussed on serious games more broadly. Reviewing the role of inoculation theory is particularly important as these inoculation serious games differ significantly to conventional inoculation messages in length, content, and the context in which they may be engaged with. A review of game-based inoculation interventions not only draws a picture of the current state of this research but can guide practitioners creating new interventions. Through completing this review, this chapter has contributed to addressing the third research question: how can game-based inoculation interventions against misinformation be better designed.

25 serious games have been reviewed on game design (i.e., topic, genre, platform, target audience and roleplaying perspective) and inoculation theory variables

(i.e., threat type, counterarguments type, and existence of components such as a forewarning). This review goes further than reviewing game-based interventions by also reviewing the transparency and discussion surrounding the design of the interventions in corresponding papers, including formal design processes, design considerations, and interaction with (serious) game design literature.

This chapter offers three contributions to the literature. Firstly, this review finds that although scholars have engaged with a wide range of real-world issues, interventions have relied on largely digital platforms following a similar genre. The potential that other platforms such as analogue design and virtual reality may have to not only differently elicit inoculation mechanisms, but support learning is underexplored. A pattern of precedent has also emerged in game genre, with many designers adapting previously successful interventions rather than exploring new designs. As the previous chapters of this thesis have indicated, different platforms and designs such as analogue and collaborative designs may have the potential to last longer and better support learning.

Secondly, a critical review of how game-based interventions elicit inoculation theory mechanisms revealed a lack of clarity and transparency of exactly how interventions' design achieved this. I find that, despite the recognition of inoculation theory as a foundational theory in many games, there is often superficial discussion describing how the theory has informed design. For example, of the 25 games reviewed, only 6 have corresponding papers that state how threat – the central mechanism to inoculation theory (Compton, 2024) – is elicited (i.e., through a forewarning, or implicitly through game content). A closer review of the design of the reviewed games casts further doubts on the efficacy of intervention design to elicit this key mechanism and reveals that design decisions are even at times contradictory in their justification.

Finally, this review finds an under-discussion of game design processes and argues that more transparency and interaction with (serious) game design literature surrounding design processes may help practitioners create more effective interventions. Game design literature has long explored game design processes such as 'playcentric' design (see Fullerton, 2004) to support the development of games that are engaging and enjoyable. These processes have been successfully applied to serious games, for which established instructive frameworks now exist (Culyba, 2018). More transparency and discussion of game design processes undertaken, interaction with game design literature, and justification of design decisions are recommended, as this can facilitate richer comparison between approaches, and guide practitioners to make more effective interventions.

This review has highlighted a wide range of promising research with game-based inoculation interventions against misinformation. Many game-based interventions have successfully improved individuals' resilience to misinformation on important topics and has reached large audiences around the world. Nonetheless, this review has identified that the field is in fact in its infancy, with many game genres, platforms, and methodological approaches yet to be comprehensively explored. Importantly, this review finds that many findings from Chapters 4 and 5 questioning the role of inoculation theory and highlighting the importance of game design apply to the wider game-based inoculation field. However, this chapter is not complete in answering research question 3. This chapter has highlighted how current game-based approaches are currently designed and highlighted areas for further investigation, but importantly does not explore any formal, instructive approaches. The next chapter in this thesis aims to apply many of the findings from this review by designing a new game-based intervention.

7. The disPHISHinformation Game: Creating a Serious Game to Fight Phishing Using Blended Design Approaches

Parts of this chapter are taken from publications by the author of this thesis, 'The disPHISHinformation Game: Creating a Serious Game to Fight Phishing Using Blended Design Approaches', in Human Factors in Cybersecurity AHFE Proceedings.

The previous chapter has reviewed game-based interventions targeting misinformation, and identified areas where future interventions could improve, particularly on their theoretical grounding, exploration of game platforms and genres, and transparency of design decisions and processes. The previous chapter does not insist on specific intervention designs that practitioners should follow; instead, the review recommends and practitioners to explore new and novel game-based intervention designs. This chapter is intended to begin to provide instructional support for game-based intervention designers by demonstrating a theory-led application of game-based inoculation to a novel issue: the human aspects of cyber security. By drawing on findings from the previous three chapters, and through the development of a new game-based intervention, this chapter aims to contribute to answering the third research question:

> **RQ3**. How could game-based inoculation interventions against misinformation be better designed?

Through collaboration with the large multinational insurance company AVIVA, this chapter presents the serious game *The disPHISHinformation Game*, a game-based intervention to increase players' resistance to an important type of misinformation: digital social engineering. The intervention is an analogue intervention with collaborative elements, where players must decide the veracity (i.e., real or malicious) of incoming messages. The intervention is based on real examples of digital social engineering attacks, whilst drawing on the (serious) game design and development literatures. In the game, players must decide if messages are genuine or malicious (a phishing, vishing, or smishing attack); on the rear of each card is written

the answer, plus techniques players could use to determine the message's veracity in the future. Action cards enable players to also work collaboratively to determine whether a message is real or malicious, allowing a player to ask a person/people to assist them. Pilot results from a playtesting session informed further extension to collaborative gameplay features.

Informed by the previous chapters of this thesis, this chapter does more than just introduce the intervention. Instead, care is taken to a) discuss the theoretical grounding of the intervention, b) highlight the instructive design frameworks used to develop the serious game, and c) showcase important preliminary findings from a pilot playtesting session in an organisational context. Promising findings from Chapter 4 and 5 on the value of tangible embodied game elements afforded by analogue designs has informed the genre of the intervention. Findings from the review in Chapter 6 have highlighted the importance of a considered theoretical grounding and discussion around design decisions and process. The focus of this chapter is not to measure the efficacy of the intervention; instead, this chapter's contribution is discussion of the development of *The disPHISHinformation Game*, with particular effort to highlight the design approach taken.

This chapter is structured in four sections. First, the chapter gives a comprehensive overview of the intervention itself, introducing the game materials and gameplay characteristics. The second section establishes the theoretical grounding for the intervention, with the chapter drawing on established inoculation message design protocols to outline how the textual content and format for the game-based intervention has been created. The third section documents the design processes followed to create the intervention, drawing on established serious game design and development processes, namely, The Transformational Framework and the 'playcentric' game development process. Building upon the intervention presented in Section 7.1, the final section of this chapter presents findings from a small-scale public 'playtest' session with the intervention at the partner organisation, AVIVA. This section not only demonstrates the value of following established design processes but showcases how findings from this public playtest have further informed the design of *The disPHISHinformation Game* itself.

7.1 *The disPHISHinformation Game*: An Overview

To give context to the reader this short section gives a comprehensive overview of *The disPHISHinformation Game*, introducing the intervention's rules, cards, and mechanics. The version of the intervention presented in this section precedes the public playtest discussed in the final section of this chapter. First this section introduces the three

different types of cards used in the intervention: the context card, source cards, and action cards. Following this, the specific mechanics of play are descried.

The first card used in the intervention is a **context** card. This card serves no purpose in gameplay; instead, it gives player contextual information and delivers a forewarning (see Section 7.2). When receiving a malicious digital social engineering attack message, the context of topic, personal information, and indirectly associated information (e.g., workplace) communicated are important to help identify it as malicious (UK National Cyber Security Centre, 2017). In *The disPHISHinformation Game*, players play as employees of the fictitious company 'Creative Ads', working within a project management department (see Section 7.3 for playtest-informed design decisions on this). Players are also informed that Creative Ads uses the creativeads.co.uk domain. This added context enables the inoculation of more context-specific digital social engineering identification techniques (e.g., hostname obfuscation). This context card (see Figure 24) also contains a short forewarning message alerting the player that attacks are likely (the longer forewarning message defined in Section 7.2 was not implemented in the intervention until after the public playtest outlined in the next section).

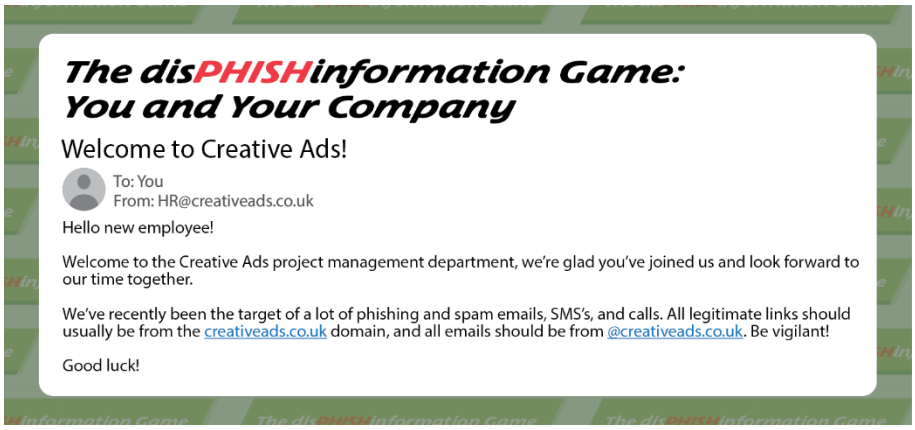


Figure 24: The game context card and forewarning (www.disphishinformation.org).

The core part of *The disPHISHinformation Game* are **source** cards, which represent messages players receive while working at 'Creative Ads'. The game's design is centred around digital social engineering attacks that are sent to players. A variety of different characteristics and manipulative techniques (see following section) are featured in source cards, imitating phishing, vishing, and smishing attacks commonly observed by organisations. Half of the source cards (50 cards) are 'weakened dose' attack message, with detailed refutations (see Section 7.2) on the rear detailing how the player could have identified it as malicious. The other half (50 cards) are 'safe' source cards containing legitimate organisation messages; the cards share an identical design on the front, with more details on the rear of cards. Attack source cards have two to three refutational messages on the rear, each refutation specific to the type of attack it is (e.g.,

the technique used to motivate user engagement with the message, the type of attack, and any obfuscation methods used, see from Section 7.2 Table 13). Example of some attack and refutation messages are available in Appendix 11. See Figure 25 for example 'safe' and 'phish' source cards.

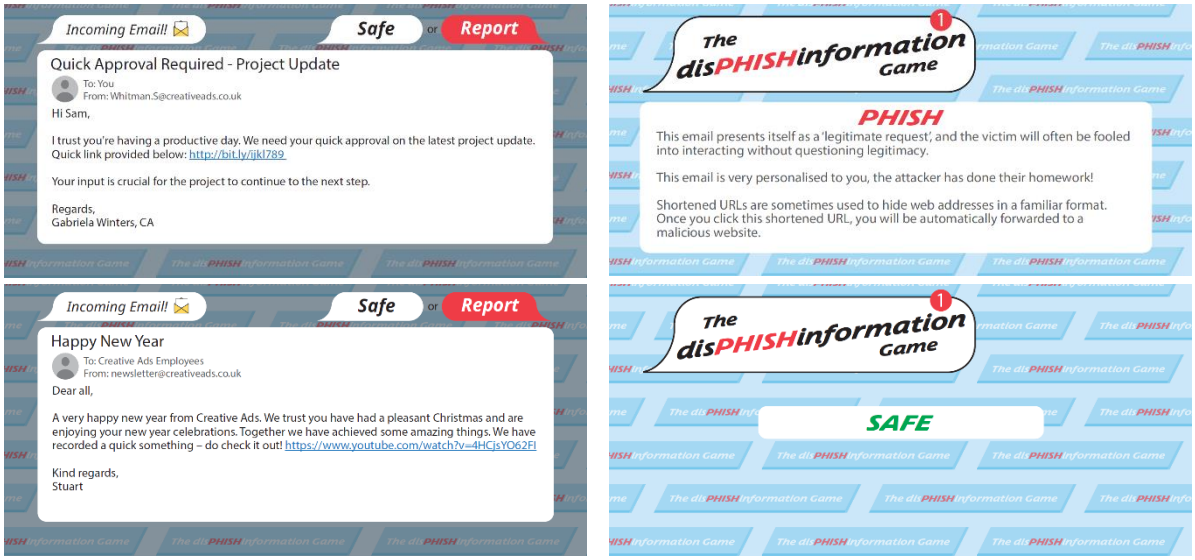


Figure 25: Source cards of the disPHISHinformation game. Shown above is message content (left) and (on rear of card, right) if that message is malicious, or real.

The disPHISHinformation Game uses **action** cards to facilitate collaboration and collaborative learning between players. Action cards allow players to either a) ask another player to help them decide the veracity of a source card, b) ask all other players to help them decide the veracity of a source card, c) pass the source card to another player, or d) skip their turn (see Figure 26). Action cards have been designed for two purposes, to foster humorous interactions between players (maintaining player engagement/retention and increasing chances of replay), and allow players to learn from one another. When a player must determine the veracity of a difficult source card, they may use action cards to collaboratively learn from other players around the table. Earlier chapters of this thesis observed promising moments of collaboration when players completed the game-based intervention *Fake News*. Integrating collaborative learning elements into the design of this intervention allows players not only to be inoculated through game materials, but to learn identification techniques from one another.

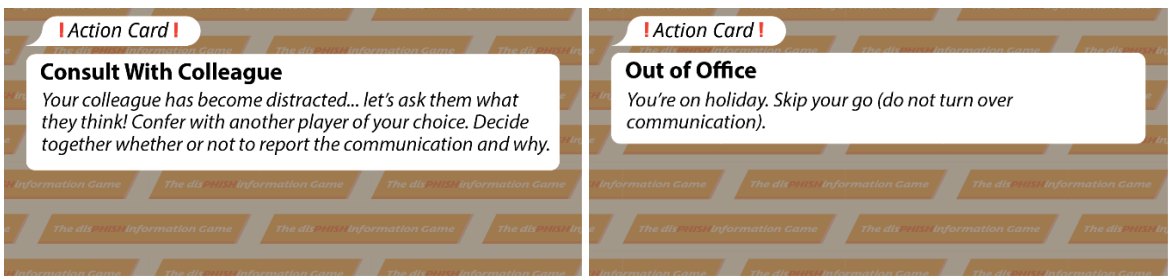


Figure 26: Two examples of disPHISHinformation action cards.

Before the game starts, Source and Action cards are shuffled together into a single deck. The deck is then placed between all players such that the content of the source/action card faces up, and the source card answers (real or 'phish') face down. Players are each given a context card for reference which they keep during play (this card is non-playable). On a player's turn, the player reads the card at the top of the deck. If it is an action card, they can take this card and keep it for future use. If it is a source card, the player must decide if the incoming message on the source card is genuine or malicious. Once they have come to a decision, they declare this to the group and turn the card over. If correct, the player gains a point. If the source card is 'phish' (an attack message), the player must read aloud the attack refutations on the rear of the card so not only the current player is inoculated, but all players can gain benefit from the card. Players continue until a player reaches a certain number of points (this can be adjusted depending on available time or context).

Despite the existence of cyber security serious games from industry and academia, many large-scale organisations either continue to use conventional training approaches or develop gamified approaches in-house. As well as a natural resistance to adopting experimental approaches, medium and large-scale enterprises typically have security policies specific to their organisation. This requirement for organisation-specific training content can reduce appetite for adopting serious games that have no flexibility to be customised. *The disPHISHinformation Game* has been designed with customisability in mind for this reason. Blank source cards have been provided alongside the original content, and users are able to edit these and use content more specific to their organisation.

This section has given an overview to the intervention *The disPHISHinformation Game,* presenting the three different types of game cards (context, source, and action), and the intended gameplay. The design of this intervention has been informed from not only inoculation and learning theories (see Section 7.2), but instructive design and development processes such as serious game design literature and 'playcentric' prototyping. The final section of this chapter offers findings from the presented intervention's first public playtest, outlining observed issues and future design changes.

7.2 Designing Textual Format and Content: Theoretical Grounding

A large proportion of serious games against misinformation have cited inoculation theory as a leading concept informing their design (Kiili et al., 2024). Many of these interventions have been effective in not only improving the veracity discernment of information but meeting other objectives such as attitudinal or behavioural changes (e.g., building resistance to extremist content and reducing vaccine hesitancy in players,

Hopkins et al., 2023; Saleh et al., 2023, 2024). The previous chapter of this thesis has more closely explored game-based interventions and their grounding in inoculation theory. An important finding from this review is that many of these interventions draw on inoculation superficially: drawing from the inoculation analogy as a pre-emptive approach to fighting misinformation but not interacting with the theory's mechanisms. I have sought for *The disPHISHinformation Game* to have a stronger theoretical grounding, drawing on not only inoculation theory but collaborative learning theories as well. This section presents the theoretical grounding informing the design of the intervention, not only justifying the design of the intervention, but demonstrating how inoculation theory can be drawn on in game-based contexts.

Early work by McGuire was limited to inoculating individuals against attacks on widely accepted beliefs on single issues, such as teeth-brushing and the benefits of penicillin (McGuire & Papageorgis, 1961). In these contexts, refutational pre-emptive message components introduced and refuted arguments specific to these topics (see Section 2.2). Only more recently have scholars explored whether technique-based inoculation (i.e., inoculating individuals on misinformation techniques rather than specific arguments) can allow inoculation to be scaled to mitigate misinformation more broadly. Technique-based inoculation has been found to be effective (Roozenbeek et al., 2022; Roozenbeek & van der Linden, 2020; Saleh et al., 2024), but scholars have found that in side-by-side comparisons issue-based interventions can in fact outperform technique-based inoculation (Banas & Miller, 2013), and is especially effective when one knows the misinformation individuals may be exposed to later (Roozenbeek et al., 2023; Zerback et al., 2021). Crucially, issue-based research has found that counterargument refutations do not actually need to cover all potential attack arguments on a topic; resistance has been found to transfer to other arguments as well to provide 'blanket protection' within the issue domain (Banas & Rains, 2010; Ivanov et al., 2009; Parker et al., 2012, 2016). Moreover, issue-based inoculation has greater established instructional support from inoculation literature, with much research exploring message and component design and key mechanisms (Compton, 2012, 2024).

This novel intervention addresses a type of misinformation with real world impacts, digital social engineering (e.g., Badshah, 2025; Warren, 2025). Cyber security is a key focus for organisations around the world, with 43% of UK businesses reporting having experienced a cyber security incident in 2024 (Rizvi & Fordham, 2025). For organisations, cyber security attacks risk affecting the confidentiality, availability, and integrity of data (Dhillon & Backhouse, 2001), impacting company finances and customer trust, and in the case of cyber-physical systems, threatening human life (Oman et al., 2002; Ralston et al., 2007). For governments, cyber security attacks can affect sovereign structures and even influence democratic processes (Rid & Buchanan,

2018). Digital social engineering attacks have remained the leading vulnerability in organisations' security posture; in 2024 84% of reported cyber security attacks on UK businesses were phishing attacks (Rizvi & Fordham, 2025). The human aspects of cyber security have become an area of great interest (Ghadge et al., 2020; Ulven & Wangen, 2021); through education, cyber security skills can be taught that help employees defend both themselves and their organisations from continuously evolving attacks. Cyber security training, often with a focus on phishing, is now becoming more common in organisations (Coutinho et al., 2023). Conventional training can include point-and-click exercises, watching video media, or simulations. Despite phishing training being commonplace in organisations, the risks of a successful attack have remained high (Coutinho et al., 2023; Rizvi & Fordham, 2025) and some scholars have raised doubts on the effectiveness of current phishing training practices (Fernando & Arachchilage, 2019).

As well as being a pressing issue with serious real-world impacts, the human factors of cyber security may be a particularly appropriate application of inoculation theory. Although much recent inoculation scholarship has focussed on misinformation, the theory was incepted to protect individuals from persuasion more broadly and has in fact been applied to a wide range of topics (Compton, 2024). Research into phishing attacks indicates that the psychological processes at play share many similarities to misinformation: both use psychological manipulation techniques (Bago et al., 2020; Ecker et al., 2022; Sarno et al., 2020, 2021; Sarno & Black, 2024; Swami et al., 2014; Wardle & Derakhshan, 2017) such as emotional content (Brady et al., 2020; Roy et al., 2025) to bypass rational decision-making. Inoculation theory may therefore be a particularly suitable mitigation strategy to reduce individual vulnerability to digital social engineering attacks, and some promising research has already shown that inoculation-driven phishing interventions can be effective (Roy et al., 2025; Wu et al., 2020). Social engineering (i.e., user manipulation to facilitate a cyber security attack) is not a specific narrative issue on which there are arguments 'for' and 'against', such as conventional issue-based interventions (e.g., climate change, vaccine hesitancy). Nonetheless, refutations of example social engineering attack messages (e.g., phishing emails) may transfer to other attacks within the issue-domain, providing a 'blanket protection' (see Parker et al., 2012, 2016) for players from digital social engineering attacks. Given this, and the instructional support provided by issue-based inoculation literature, an issue-based inoculation approach has been adopted in the design of this intervention.

Scholars have already explored the use of analogue serious games to 'teach' players on and raise awareness of the dangers of phishing. *Riskio* (Hart et al., 2020) aims to build knowledge in non-technical players within an organisational context. One

player acts as the attacker with other players acting as defenders, working collaboratively to pick appropriate 'defence cards'. *Decisions & Disruptions* (Frey et al., 2019) educates players on cyber security decision making and prioritisation. Players play as security decision-makers for a utility company over 4 rounds. Each round, players decide what security defence to invest in with their limited budget. A Lego game board represents the players' facility and office; however, the author has previously played *Decisions & Disruptions* solely with cards with little issue. *Decisions & Disruptions* has become particularly popular within cyber security circles, as is currently deployed as a training tool by the Metropolitan Police's Cyber Protect team (Metropolitan Police, 2023).

Designing Inoculation Message Components

Following an issue-based intervention design has allowed me to draw from more instructional elements of inoculation theory literature, particularly conventional forewarning and pre-emptive refutational task components. Although not required, a forewarning message component can provide an explicit form of threat to the individual (Compton, 2024; McGuire & Papageorgis, 1962). A forewarning has been integrated into the design of *The disPHISHinformation Game* and is provided for players to read alongside gameplay instructions. Forewarning design has followed established inoculation message design protocol (Ivanov, 2017; Parker et al., 2012; Pfau et al., 2010; van der Linden, 2022). In particular, the forewarning component for *The disPHISHinformation Game* has been designed to achieve established forewarning design techniques by (1) highlighting to players that attacks are not only possible but likely (Ivanov, 2017; Parker et al., 2012), (2) informing players of their vulnerability to attacks (Parker et al., 2012, 2025) (3) communicating that other individuals who felt well-prepared have frequently succumb to social engineering attacks (Ivanov, 2017; Parker et al., 2012), and (4) emphasising the impact phishing attacks have on not only businesses and organisations but individuals as well. The forewarning provides evidence for all claims made (communicating factual evidence for all claims made is an important strategy, see Ivanov, 2017; van der Linden, 2022), and emphasises the severity and personal significance of the inoculation issue (see Pfau et al., 2010). The final forewarning can be seen in below (note that this forewarning message was only implemented after the public playtest of the intervention, as outlined in Section 7.4).

> "Phishing attacks are by far the most common form of cyber-crime, with ~3.4 billion spam emails being sent every day. In 2024, 84% of reported cyber security attacks on UK businesses were phishing attacks. A 2019 Threat Report by Symantec showed that 65% of cyber-attacks were executed through spear phishing. A single successful phishing attack can have a disastrous impact on not only organisations (e.g., through loss of

> data, impact on reputation, and regulatory fines), but individuals as well. Research shows that people are overconfident in identifying phishing: you may not be as good at spotting phishing messages as you think you are." – *The disPHISHinformation Game* forewarning message

The second component in conventional issue-based inoculation design is a refutational pre-emption task. This component raises and refutes false arguments on the issue with the individual being either passively or actively involved in this activity. This message component not only equips individuals with counterarguments to future persuasive attacks, but can implicitly increase threat, and even improve counterarguing skills (Compton, 2024). It is with the pre-emptive refutation component that this intervention breaks with conventional inoculation message design. Conventionally only two or three relevant oppositional arguments are often used (Ivanov, 2017). Although some specific phishing scams have gained notoriety (e.g., the 'advance-fee' scam, see Glickman, 2005), most are varied in design and technique. Given that it is nearly impossible to inoculate against every digital social engineering attack message players may receive individually, this intervention must rely on the blanket protection ('umbrella' protection, see Parker et al., 2016) afforded from an inoculation intervention (i.e., for resistance conferred through refuted arguments to also protect against novel arguments). The pre-emptive refutational messages used in this intervention draw from memory literature and the importance of repetition on learning (Wimmer & Poldrack, 2021); the pre-emptive refutation component includes more (but shorter) oppositional arguments to reinforce learned content from refutations. In practice, the inoculation content for this intervention contains a large number of example digital social engineering attacks, each of which is not only refuted, but refuted with two to four identification techniques (making reference to evidence of human vulnerability outlined in the forewarning). An example message can be seen in Table 12, but note that the design of this message has also been informed through instructive design processes outlined in the following section.

Table 12: Example 'weakened dose' attack message and accompanying refutation. The disPHISHinformation Game makes use of a large number of such attack messages to reinforce outcomes learned from refutations. Note that cards include further contextual indicators not included below (e.g., the sender's email address).

Attack message	Refutation
Hi, I hope you're well. We're conducting a brief account verification to enhance security. Please click the link below to confirm your account details: https%3A%2F%2Frandsecuritycheck.net/dl/13zf	• This email presents itself as a 'legitimate request' and the victim will often be fooled into interacting without questioning legitimacy. • The sender's email address is camouflaged to look like a legitimate email from a @createiveads.co.uk account. Can you tell the difference?

Best, Supervisor, Creative Ads	• The attacker has 'encoded' the URL to visually hide the true malicious URL address.

Chapters 4 and 5 of this thesis demonstrated that analogue intervention design may help create a memory-resistant intervention that not only incited positive affect from players and encouraged post-inoculation talk but supported collaborative and experiential learning from players. The tangible, embodied design was responsible for good memory retention and positive effect in players. As discussed in the previous chapter, although this positive affect may reduce elicited apprehensive threat from the intervention, I believe this will not negate the *motivational* threat elicited explicitly through the forewarning and implicitly in the refutational pre-emptive messages. Positive affect may also reduce inoculation hesitancy and encourage individuals to complete the intervention more than once (an important behaviour in security training, see Reinheimer et al., 2020). For these reasons (drawing on the previous study showcased in Chapter 4 and 5), the intervention is designed as an analogue card game.

The collaborative elements of the intervention *Fake News* also allowed players to learn from one another during play. Collaborative elements have therefore been incorporated in the design of *The disPHISHinformation* Game, allowing players to share skills and techniques to spot digital social engineering messages they may have. Moreover, following collaborative learning literature, I have drawn on competition to positively influence collaboration and collaborative learning (providing a social motivation for 'productivity', see Deutsch, 1949; D. W. Johnson & Johnson, 1991; P. Williams & Sheridan, 2010).

This section has outlined the theoretical grounding for the new cyber security intervention *The disPHISHinformation Game*, drawing on inoculation theory to not only elect a collaborative analogue design, but to create the inoculating content to be used in the intervention. Findings from the previous chapter highlighted that at times some game-based interventions superficially drew on inoculation theory, not interacting with the theory's mechanisms and components. This section has sought to describe in depth the theoretical grounding for *The disPHISHinformation Game*, using established instructional inoculation message design protocols to inform the design of a forewarning component and pre-emptive refutational game content to elicit threat and improve resistance to digital social engineering attacks through 'blanket protection'. Moreover, findings from Chapter 4 and 5 have informed the collaborative analogue design of the intervention.

7.3 Intervention Design and Development: A Formal Process

An important finding from the review in the previous chapter was a lack of discussion and transparency surrounding design processes. As many of the reviewed game-based interventions had superficially drawn on inoculation theory, I hypothesised that the success of these interventions may have been more closely linked to effective design processes. The first section in this chapter has more comprehensively drawn on inoculation theory design protocols, informing intervention content and an overview of how players will interact with 'weakened dose' messages (Compton, 2024). The role of design process, then, is to transform the theory-driven concepts and content from the previous section into a complete game-based intervention. Drawing on (serious) game design literature—namely The Transformational Framework (Culyba, 2018) and the 'playcentric' design processes (Fullerton, 2004)—I outline the design and development of the intervention with effort to maintain transparency of processes adopted.

The Transformational Framework (Culyba, 2018) is a requirement-gathering approach which prioritises the alignment of game objectives with broader learning or behavioural goals within serious games. This pre-production-based methodology (i.e., to be drawn on during the conceptualisation and early design stages of a serious game) which integrates explorative and reflective question-based processes for a more structured serious game development approach. A core ethos of the framework is that serious game design is highly interdisciplinary (e.g., game designers, artists, educators, and domain subject experts all collaborate in the creation of a serious game), and organising critical considerations from these different disciplines can help serious game practitioners (particularly in small teams) to make considered design choices. The framework itself has eight pieces and, although all are important, this section will focus on domain concepts, barriers, and audience and context (other parts of the framework, such as 'high-level purpose', 'player transformations' and 'prior works' have already been covered in the previous section of this chapter).

The 'domain concepts' tenet of The Transformational Framework considers the subject-specific information communicated to players (i.e., the content in the serious game to 'transform' players). Importantly, the framework presses practitioners to adopt not only a broader view of intervention outcomes (see 'types of transformations', Culyba, 2018, p. 90), but a broader characterisation of domain concepts, going beyond communicating simple facts and considering semiotic, procedural, explanatory, narrative, affective, and practical player 'transformations' as well. This aligns well with findings from Chapter 5, where the learned outcomes of the intervention were found to be broader than just behavioural transformations in players.

However, the framework also stresses the importance of a tightly defined scope to ensure a manageable development process. The previous section of this chapter has outlined (drawing from inoculation theory) how players must interact with (weakened) social engineering attack messages. The scope of domain concepts in *The disPHISHinformation Game* are limited to different phishing, vishing, and smishing characteristics and manipulative techniques (see Table 13) collated from cyber security literature (Alabdan, 2020; Aleroud & Zhou, 2017; Rastenis et al., 2020).

Table 13: Catalogue of digital social engineering attack features from which intervention content is drawn from.

Attack Platforms	Engagement Motivation	Sender's Email/Number	Attack Type
Phishing	Benefit for recipient	Owned	Behavioural
Smishing	"Legitimate" request	Camouflaged	Shortened URL
Vishing	Important information	Hacked legitimate	Bad domain name
	Possible failure		Hostname obfuscation
			Encoded URL obfuscation
			Malicious attachment

Pre-emptive refutational message design requires the 'weakened dose' social engineering attacks to also be refuted. Established inoculation message design protocol discussed in the previous section has outlined that digital social engineering attacks should be refuted with two to four identification techniques (e.g., Ivanov, 2017; Parker et al., 2012, 2016). The 'domain concepts' tenet of The Transformational Framework (Culyba, 2018) provides more instructional support on the types of concepts that can be communicated to players. Refutations to phishing, vishing, and smishing attacks in the present intervention have been designed to not only give factual information but cause-and-effect explanations, step-by-step rules, and identification techniques. Naturally, many refutations used use more than one type of domain concept. Some example message refutations are listed in Table 14.

Table 14: Example refutations to 'weakened dose' attack messages used in The disPHISHinformation Game.

Domain concept type	Example refutation used in intervention
Factual	This phishing email has come from a legitimate Creative Ads account that has been hacked.
Procedural and semiotic	The attacker has included a malicious attachment to this email. Attachments that are .rar, .exe, .zip, and even .docx and .pdf can be malicious. Always check first.
Procedural	The attacker has used a domain that looks very similar to www.creativeads.co.uk. Can you spot the difference?

Explanations	This call gives the impression of an important event, and the stress caused often overrides the victims sense of phishing security!

The Transformational Framework also draws attention to the 'barriers' preventing or discouraging players being positively impacted (i.e., 'transformed' or educated) by the intervention. Something that has received attention in inoculation literature is inoculation hesitancy: exploring how to make an intervention more attractive to encourage its organic uptake (A. Johnson & Madsen, 2024). The uptake for *The disPHISHinformation Game* has been considered through two design decisions. Firstly, an analogue game-based inoculation format has been selected. Previous findings in this thesis showed that analogue game design was particularly popular with the adult sample, and this demographic trend is supported by broader literature as well (Cross et al., 2023). This is important as this is a demographic that has been found to be particularly vulnerable to phishing attacks (Sarno et al., 2020). Secondly, the intervention has been designed with the manufacturing requirements for organisations in mind. Many existing (although effective) cyber security serious games use elaborate game pieces made from plastic, causing high costs for organisations hoping to adopt these interventions as part of cyber security training policies (see Denning et al., 2013; Gondree & Peterson, 2013; Graffer et al., 2015). By adopting a purely card-based design, this intervention can be both cheap and accessible for organisations who can print materials at little cost.

Finally, The Transformational Framework encourages practitioners to consider not only the target audience for the intervention but also the context in which it is designed to be played. These key considerations can impact the design of an intervention, and findings from the previous chapter found these variables to be underdiscussed in many current game-based interventions. Although some demographics have been shown to be more vulnerable to phishing attacks than others (Sarno et al., 2020, 2021), this intervention does not target a specific audience. Nonetheless, two design considerations have been made in relation to audience. First, perceived complexity of an intervention has been shown to discourage its engagement and uptake (Zagal et al., 2006). The interventions should therefore be accessible to not only younger players but older players who may struggle to pick up rules as intuitively as regular players of (analogue) games. Secondly, many organisations face cyber security threats (i.e., phishing techniques) that are specific to their sector, product, or company. *The disPHISHinformation Game* allows organisations to customise the intervention to meet the specific risks unique to them by providing blank cards for them to write their own 'weakened dose' phishing attacks and refutations.

The Transformational Framework has influenced the development of this game-based intervention in other dimensions as well and has the potential to provide

even further instructional support to serious games that may not have the external theoretical grounding that this intervention does. The four key design considerations defined from the discussion above are listed below for reference. The next step in the creation of a new game is development and development processes (Fullerton, 2004).

C1. To appeal to a wide socio-demographic range (particularly individuals more susceptible to attack), the game-based intervention will have an analogue design.

C2. All individuals are vulnerable to attacks, not just 'gamers'. The intervention must be designed with simple and intuitive rules to maintain accessibility for all players.

C3. To ensure low cost and high accessibility for organisations (through whom intervention uptake may be encouraged with employee security training), the game-based intervention will be made from card and be simple to print.

C4. Organisations often face phishing attacks that are not just sector-specific but can be specific to the company itself. The intervention must allow for some customisability so organisations can create their own resources.

The 'playcentric' game development philosophy (Fullerton, 2004) dictates that consideration of player experience throughout the development process (i.e., from conceptualisation to completion) is essential to creating effective games. Central to this is prototyping and playtesting. Through continuous iterative development of prototypes, playtesting of prototypes, and revisions to design based on this feedback, problems in gameplay can be caught early in the development processes, and games can be refined to "reach their full potential" (Fullerton, 2004, p. 12). Fullerton stresses that this iterative process should happen as soon as possible in the game design process, as design flaws are far more costly in time, money, and effort when addressed later during development (see Figure 27 for a visual model of this). The importance of continuous iterative testing is also emphasised in serious game design literature: "the first version of any game or game system is often confusing, broken, and generally not fun. It is only through iteration that great games can be developed" (Culyba, 2018, p. 31). A 'playcentric' approach has therefore been adopted to develop the content and considerations outlined in this and the previous section into a working game-based intervention.

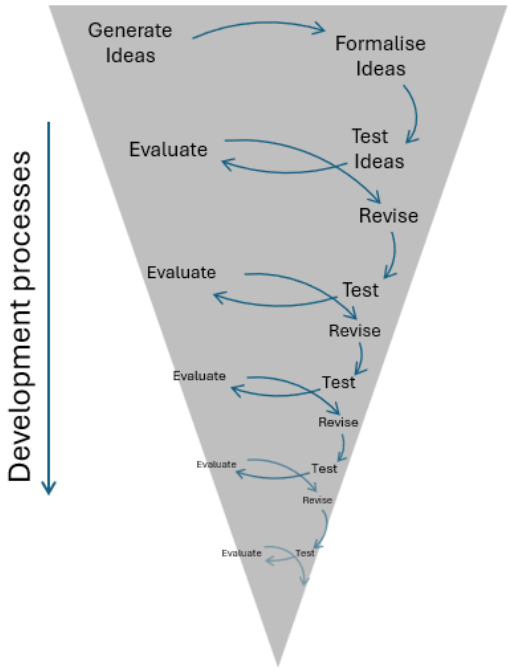
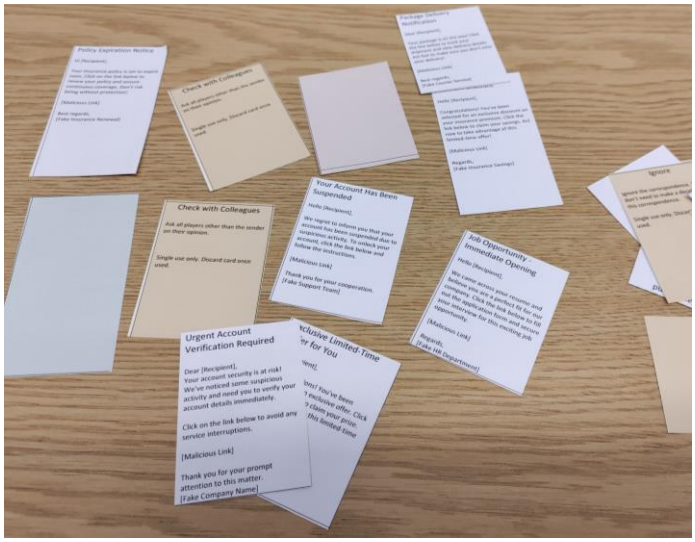


Figure 27: Fullerton's 'playcentric' iterative design throughout the production cycle. Not shown: production cycle stages.

The rest of this section overviews the prototyping and self-playtesting completed in the development of *The disPHISHinformation Game* (an overview of a public playtesting session is available in Section 7.4). We note that not all of the design considerations outlined in this (and the previous) section have been integrated into the early prototypes of the intervention. In fact, some elements such as the forewarning message described in Section 7.2 was not introduced in the intervention until after a larger playtesting session presented in Section 6.4.

Following game design protocol (Fullerton, 2004; Salen & Zimmerman, 2004; Schell, 2019), the initial game prototype was created on paper (see Figure 28). This early prototype included the core mechanics of the game-based intervention: 'weakened dose' attack messages and refutations. Constricted to the small card size these social engineering attack messages were short, and refutations were simple (colour on rear denoting attack message veracity). Self-playtesting determined that the size and orientation of the cards limited the 'weakened dose' attack messages to being too short in length (approximately one to two sentences) and not representative enough of real-world attack messages. Many contextual indicators of phishing, vishing, and smishing message veracity can only be identified in longer textual content (Alabdan, 2020; Chiew et al., 2018). Refutations in this early prototype only stated the veracity of the attack message as being 'phish' or 'safe'; this was determined to be too simplistic and, as explored in Section 7.2 of this chapter, did not follow established inoculation message design protocol. Later prototypes followed the outlined message design from section 7.2. Playtesting also determined that attack message veracity can also be determined

from contextual indicators such as the sender's knowledge of relevant situational or personal information (e.g., personal or company name), or the following of business norms in message tone and structure (Alabdan, 2020). This early prototype could not inoculate on these important identification techniques.



Package Delivery Notification

Dear [Recipient],

Your package is on the way! Click the link below to track your shipment and view delivery details. Act fast to make sure you don't miss your delivery!

[link]

Best regards,
[Courier Service]

Figure 28: Example cards from prototype 1 of The disPHISHinformation Game. On right: example card with some details yet to be determined.

The second prototype of *The disPHISHinformation Game* improved on many of the highlighted issues with the initial design. First, cards were enlarged and oriented horizontally to accommodate more detailed attack messages. Moreover, refutations on the rear attack message cards more closely followed design protocol outlined in the previous section of this chapter (see Figure 29). This prototype also introduced narrative into the intervention: players played as employees of a fictitious advertising company 'Creative Labs'. This allowed attack messages to be more detailed and organisation related. Nonetheless, some issues still remained. Firstly, the text on cards (including refutations) were still too small and difficult to read. The cards used in the released version of the intervention were larger with larger easier-to-read text. Secondly, it was found that when playing with a larger number of players there were not enough 'weakened dose' cards. Prototype one and two contained 16 attack cards (and 16 'safe' cards); the final version of *The disPHISHinformation Game* (see Section 7.1) used 50 attack cards (and 50 'safe' cards) in total.

Figure 29: Example cards from prototype 2 of The disPHISHinformation Game. On right: example card attack and refutation on rear.

This section has outlined the design and development processes of the game-based intervention *The disPHISHinformation Game*, informed by instructive frameworks such as The Transformational Framework (Culyba, 2018), and the 'playcentric' development process (Fullerton, 2004). First, this section has outlined how The Transformational Framework has (with key content established in Section 2 of this chapter) guided the design of the intervention, creating not only the content for the intervention but informing the definition of four key design considerations. Second, this section outlined the 'playcentric' development approach adopted throughout the development processes for the intervention. Two key stages of early prototyping are presented, with particular attention brought to the issues identified at each prototype stage and how these have been addressed. The review in the previous chapter found little transparency and discussion surrounding design considerations and processes in the reviewed interventions. This section attempts to maintain transparency surrounding the factors that have informed the design of the intervention itself, to not only create a more effective intervention, but to help guide practitioners and scholars to develop more effective game-based intervention design processes.

7.4 Public Playtest in an Organisation Context

Playtesting (i.e., trialling a prototype of a game currently in development with a small group of people) is an essential step in game development, and a process that has underpinned the development of the intervention presented in this chapter. Section 7.3 of this chapter has presented results from two self-playtesting sessions (a playtest completed by the designer themselves), resulting in key design and content changes to the intervention. A 'playcentric' approach, however, asks that designers involve *players*

in the development processes as soon as possible to ensure necessary changes are easily integrated without causing greater impact on finances, time, or effort later on. This chapter presents findings from the first public playtest with *The disPHISHinformation Game*. The aim of this chapter is not to ascertain the efficacy of the intervention but instead highlight areas where the design of the intervention may be improved for greater entertainment and 'player transformations' (Culyba, 2018).

Playtesting was completed at AVIVA[17], a large multinational insurance company with an office located in Norwich, UK. As a large corporation, AVIVA faces a large number of cyber security threats from a wide variety of sources, and employee security training is an essential process in protecting both the organisation itself, and employees personally. The playtest (N = 13) was mixed methods in nature, with players completing parts of a survey before and after completing the intervention, and informal observations were carried out by the researcher during play. Some elements such as questions surrounding awareness of cyber security threats and known phishing techniques were repeated in the survey both pre- and post-test. Others, such as socio-demographical information (e.g., age, education, nationality) and Likert-scale and open-ended questions on familiarity with games and technology were only completed once. The post-test portion of the survey also included open-ended questions on the intervention itself, such as clarity of instructions, elements that the player enjoyed or did not enjoy, and any additions they would like to have seen in the intervention. Following an instructive guide from The Transformational Framework the survey was designed to not only measure enjoyment and feedback, but educational/cognitive benefit from the intervention itself (Culyba, 2018; Roozenbeek et al., 2024).

Results pertaining to players' enjoyment of the intervention was overwhelmingly positive, with all players responding between 3 and 5 on a Likert-scale between 1 and 5 (1 = *not at all* to 5 = *very much*, M=4.43, SD=1.05). Players were asked to self-report their 'cyber-awareness' both before and after taking part in the intervention (5= *Very cyber aware and always on the lookout for phishing*, 1=*Absolutely not cyber-aware with little understanding of phishing*). A two-tailed paired T-Test discovered that the intervention did not significantly improve player's self-reported cyber awareness (t(13) = -0.43, p = .67). This was surprising as, when asked how much new information they felt they had learned from taking part, responses were positive (where 1 = *nothing at all* to 5 = *huge amount*, M=3.29, SD=1.14). This indicates that players believe they may have learned digital social engineering identification techniques from the intervention, but do not feel a higher level of alertness or awareness. This suggests that the intervention

[17] https://www.aviva.co.uk/

is successfully educating players but did not increase player recognition/awareness of attitudinal vulnerability, a core inoculation mechanism closely linked to alertness.

Qualitative findings to open-ended survey questions, however, demonstrated that many of the considerations made during design (see Section 7.3) had been successful. One was that the intervention should be simple and intuitive to play for all players, even those who do not frequently play games. Of the players that self-reported playing games in their personal life 'infrequently' or 'very infrequently', all reported that the instructions to the intervention were clear: "Instructions were clear and easy/intuitive to start [the] game" (playtester 6). The variety in attack messages (a result of adopting a broad characterisation of 'domain concepts', see Section 6.2) were also well represented in survey responses, with multiple playtesters (playtesters 1, 4, 6, 8, 14) explicitly highlighting this as a positive aspect not always seen in traditional security training: "this [game-based intervention] was more enjoyable and included a greater range of scenarios" (playtester 1). Similarly to previous research presented in this thesis (see Chapter *5*), the collaborative elements were highlighted as improving engagement and enjoyment: "playing with colleagues does make you focus more so you are inclined to learn more" (playtester *5*); "shared experience is more memorable than PowerPoint quizzes" (playtester 10).

The design of the intervention has drawn on collaborative learning to facilitate inoculation not just from game materials but from other players as well. However, informal playtest observations found a lack of engagement with collaborative elements of the intervention. First, players rarely used action cards to collaborate with other players as a) there was no penalty for getting an answer incorrect, and b) players did not want to 'gift' other players points: "Action cards don't really give incentive to use. There needs to be an element of jeopardy" (playtester 8). Second, while it was not a player's turn, they were frequently observed disengaging from the intervention, not listening to the attack messages and refutations read aloud by the current player. These findings share many characteristics with collaborative game design 'pitfalls' outlined by Zagal et al. (2006): "for a game to be engaging, players need to care about the outcome and that outcome should have a satisfying result" (Zagal et al., 2006, p. 33). At present, the only incentive for a player to actively engage with the turn of another player is to learn if they were successful in gaining a point. If a player has been successful in determining the veracity of all source cards on their go, then there is little engagement incentive, as they know it is impossible for another player to score higher than them (unless they are before the player in turn order). The design of the intervention must be improved to not only improve engagement while it is not the player's current go but also improve rationale for working collaboratively with other players.

Addressing the lack of collaboration in the intervention did not require modification of existing intervention materials but the introduction of a new one. To implement a penalty for incorrectly discerning a source card's veracity (to incentivise collaboration with other players), and to make sure the outcome of a player's turn affects others (to incentivise engagement when it is not a player's turn), a new 'organisation health' card has been introduced (see Figure 30). This new card (and counter) tracks a communal health bar: when players correctly guess the veracity of a source card they gain a point, but if they are incorrect this communal health bar is reduced by one point. If this health bar reaches zero, then *all* players lose the game. This not only incentivises collaboration (players do not want to be responsible for all players losing) but increases engagement for players when it is not their turn. This communal health bar is designed to imitate how organisations are affected by cyber security attacks and, if serious enough, an attack may result in bankruptcy (e.g., see Leggett, 2025; Levy, 2025). Moreover, artwork on the card itself features modified newspaper headlines from real-world cyber security incidents, implicitly eliciting threat.

Figure 30: The 'organisation health' board, on which players track the communal health of their fictitious organisation 'Creative Labs'.

Quantitative findings from the playtesting session found that players did not self-report a higher awareness or alertness to cyber security attacks after taking part in the intervention. Although the threat mechanism in an inoculation intervention can be elicited explicitly or implicitly through pre-emptive refutations, when compared side-by-side, it is often explicit forewarnings that elicit more threat (Compton, 2024). The forewarning used in this version of the intervention (outlined in Section 7.1) only alerted players to the frequency of attacks and did not follow established inoculation message design protocol (e.g., Ivanov, 2017; Parker et al., 2012, 2016). Following the results from this playtesting session, the forewarning message outlined in Section 7.1 of this chapter was implemented into the intervention (see Figure 31). Although this playtesting session did not explicitly measure threat, future playtesting could more comprehensively measure inoculation and learning elements introduced from theory-guided intervention design.

Figure 31: Rear of 'context' card, delivering threat explicitly through a forewarning message (see Section 7.2).

This section presents preliminary findings from a playtest with the game-based intervention *The disPHISHinformation Game*. Following the 'playcentric' design processes outlined in Section 7.3, players have been brought into the development process at an early time through an initial exploratory public playtest. Through a playtest at the UK insurance firm AVIVA, players completed open-ended and Likert-scale questions on a survey before and after play, as well as informal observations during the session. This section demonstrates that playtesting plays a key role in game-based intervention development; findings informed changes to the intervention through a more established forewarning element, and a new game piece to foster engagement and collaboration.

Summary and Conclusion

Cyber security attacks pose a high risk to organisation's reputation, data, and finances. Yet, human error remains the most prominent weakness in organisations' cyber security posture. Security training for employees has been a strategy widely adopted by organisations; however, some scholars have cast doubt on their effectiveness (Fernando & Arachchilage, 2019; Reinheimer et al., 2020), and phishing attacks remain the most common type of cybercrime faced by organisations (Rizvi & Fordham, 2025). Drawing heavily on findings from previous chapters in this thesis, this chapter has sought to demonstrate how game-based interventions can be better designed, with a practical application to a cyber security context. Through the development of *The disPHISHinformation Game*, this chapter demonstrates how better engagement with (serious) game design literature (e.g., Culyba, 2018; Fullerton, 2004; Schell, 2019; Zagal et al., 2006), more comprehensive exploration of a theory grounded approach (e.g., inoculation and collaborative learning theory), and by following instructive design processes (e.g., 'playcentric' design, see Fullerton, 2004), an intervention may be better designed. This chapter has sought to contribute to answering the research question: how could game-based inoculation interventions against misinformation be better designed.

This chapter has sought to adopt a highly instructive design and development processes to create *The disPHISHinformation Game*. Findings from the previous chapter found that many game-based inoculation interventions superficially drew on inoculation theory to inform their design and lacked transparency in design and development processes. This chapter has sought to demonstrate how game-based inoculation intervention can draw on established instructive inoculation message design protocols and draw on (serious) game literature to help develop an intervention that is not only effective but enjoyable for participants. This chapter makes two core contributions to game-based inoculation literature.

First, this chapter demonstrated that it is possible for game-based interventions to comprehensively draw on established inoculation message design protocols, but the design of the intervention may have to centre heavily around 'weakened dose' attack messages and their refutation. Findings from the previous chapter highlighted that many existing game-based interventions against misinformation drew superficially on inoculation theory literature. However, this does not need to be the case. *The disPHISHinformation Game* integrates two established inoculation components: a forewarning message that is delivered in the beginning of the game, and 'weakened dose' attack messages (and accompanying refutations) around which the intervention is designed. By more comprehensively drawing on inoculation theory, this intervention can explore inoculation theory literature with more

conceptual rigour. For example, as this intervention uses refutational pre-emption components, work by Parker et al. (2012, 2016) on umbrella protection and Ivanov et al. (2009, 2016) on base attitudes can be more reliably drawn upon. The same is true with threat; as this intervention includes an explicit forewarning (and that eliciting threat has been a central concept in the design processes of the intervention), work by Compton and Ivanov (2012), Ivanov et al. (2022), and Pfau et al. (2010) on threat is more applicable. Not only can inoculation literature be more reliably drawn upon, but as this intervention shares similarities to conventional inoculation messages, future findings with this intervention may contribute to inoculation literature as well.

Second, this chapter demonstrates the potential that drawing on (serious) game design literature has to improve the quality and effectiveness of an intervention. Findings from the review in the previous chapter found that very few existing game-based interventions drew on literature that explored game design and development processes or considerations in serious game design. Moreover, findings from Chapter 4 and 5 of this thesis found that intervention design played a leading role in shaping many of the outcomes observed from the intervention *Fake News*. Drawing on serious game design literature helped define key design considerations for the intervention, some of which stimulated highly positive feedback in a public playtesting session (e.g., accessibility and intuitiveness of intervention, and the broad range of learning delivered through scenarios). Drawing on established game design processes (i.e., 'playcentric' design patterns) also helped to improve design through iterative prototyping and playtesting with players in real-world contexts. The leading role that (serious) game design literature has had in the development in this intervention raises questions for current game-based inoculation practitioners on development processes. More interaction with this literature can offer instructive guidance to practitioners when designing and developing interventions. This is key, as scholars have previously expressed frustration in not knowing 'what works' in intervention design (Appel et al., 2025; Harjani et al., 2023; Ma et al., 2023; Roozenbeek et al., 2024; Saleh et al., 2023).

This chapter has adopted a transparent approach to presenting a new cyber security game-based intervention, *The disPHISHinformation Game*. Importantly, this chapter has drawn on findings from the previous three empirical chapters on analogue collaborative game design, the role of learning theories in game-based contexts, how (serious) game design literature can inform design approaches, and the role of inoculation theory in game-based contexts. After findings from the previous chapter, this chapter has explicitly outlined the design and development approach taken in creating the new intervention. Moreover, this chapter demonstrates how inoculation theory can be more comprehensively drawn upon when applied to game-based contexts. The author hopes that through transparency of development approach,

practitioners may be guided in not only creating more effective and enjoyable game-based interventions but also exploring other development processes.

8. Conclusion and Future Work

This thesis has sought to explore how game-based individual-level interventions can improve individuals' resilience to misinformation. Specifically, the empirical chapters in this thesis have studied not only the effectiveness of analogue, collaborative game-based interventions against misinformation, but explored existing – and importantly new – design and development processes. The literature review completed at the beginning of this thesis (see Chapter 2) identified four pressing research gaps in game-based inoculation literature, namely a lack of longitudinal research, very little exploration of non-digital platforms for interventions, little interaction with pedagogy (despite a recent call to explore this in relation to inoculation theory, see Maertens et al., 2025), and little understanding of *why* interventions are (in)effective. This thesis has attempted to contribute to these research gaps following an inductive pragmatist research paradigm.

This thesis presents the findings and discussions from three studies in four empirical chapters. Each study has been designed to, at least in part, address one or more of these defined research questions. Moreover, using an inductive research philosophy has enabled a consecutive research strategy to be adopted, with findings from each chapter informing the design of the next (see Chapter 3 (methods) for further discussion on this). In this final chapter of the thesis, I aim to further contextualise and discuss the findings from these four empirical chapters. In the following paragraphs I give a top line summary overviewing the findings from these chapters. In the first section, I contextualise these findings in relation to the thesis's research questions, discussing in turn how this thesis has addressed each question. Next, I outline the thesis's core empirical and theoretical contributions. Finally, I discuss the implications of these contributions and outline an agenda for research hand practice which draws on the approach and findings of this thesis.

The first empirical chapter of this thesis presented quantitative findings from a study with the analogue game-based intervention *Fake News*. Being longitudinal and within-subjects this study represented a replication with extension of previous research with the same intervention by Roozenbeek and van der Linden (2018). Quantitative measures found no significant improvement in participants' perceived reliability or persuasiveness of 'fake news' articles, nor did participants agree with them any less. These findings represent the second failed replication of original research by

Roozenbeek and van der Linden (2018); however, the quantitative elements of this study were limited by potential floor effects. Moreover, these measures only focus on behavioural outcomes; serious game design literature highlights that outcomes from game-based interventions are often broad and can improve knowledge and skills, change beliefs, and stimulate affective responses in players (Culyba, 2018).

Following a pragmatist research paradigm, the study discussed in Chapter 4 broke with convention by also incorporating a number of qualitative methods including semi-structured interviews and session observations. Recent inoculation scholarship has proposed that game-based inoculation may be better explained by a process of learning and forgetting. Theories and models on learning have a rich history, yet scholars are yet to engage with the mechanisms of these learning models. Chapter 5 presents findings that the analogue intervention *Fake News* stimulated outcomes that can be interpreted through different learning theories, such as experiential learning, reflective learning, and collaborative learning. Previous scholars have labelled game-based inoculation interventions as "experiential learning" (Roozenbeek & van der Linden, 2020, p. 1), but this study goes further and argues that the stages of experiential learning were well represented in findings. Another learning theory, collaborative learning (D. W. Johnson & Johnson, 1987), was central to this study's findings. Session observations repeatedly found that players learned from each other as they worked collaboratively to complete the intervention. Participants were observed sharing relevant misinformation techniques with each other and applying these techniques to real world news publications and politicians. This chapter also explored participants' experience of the intervention, and the participants themselves and their experiences with misinformation. Importantly, qualitative inquiry included an interrogation of inoculation theory's role in the intervention, concluding that many of the intervention's outcomes could be more closely attributed to (serious) game design decisions and development processes than inoculation theory's influence on intervention design.

The second study of this thesis, an integrative review of game-based inoculation interventions against misinformation, sought to not only contribute to the third research question but address open questions from the previous study surrounding the role of inoculation in the design process. The review summarised the design trends of existing interventions, how inoculation mechanisms are currently being elicited through game-based intervention design, and the design and development processes adopted in these interventions. As well as identifying intervention design genres and platforms under-explored, this integrative review found that the core mechanisms of inoculation theory, namely threat, had not been comprehensively drawn upon, and the design and development processes for these

interventions were under-discussed. This chapter (Chater 6) also discusses the lack of clarity and transparency around theory-led design and development processes and emphasises that this requires greater focus in game-based inoculation scholarship.

The final empirical chapter of this study presented a new game-based intervention designed to inoculate players against a different type of misinformation, digital social engineering. This engaged exploratory research sought to more comprehensively follow established inoculation message design protocol in a game-based context. As well as a more considered forewarning component, this new intervention designed play around the 'weakened dose' refutational pre-emption component, 'gamifying' the processes of evaluating message veracity, supported by refutations. Moreover, this study sought to maintain transparency in the design and development approach adopted, drawing from (serious) game design literature such as The Transformational Framework (Culyba, 2018) and 'playcentric' development (Fullerton, 2004). This study does not outline prescriptive game-based insulation design, instead it seeks to demonstrate how to more comprehensively draw on inoculation theory and (serious) game design literature in game-based contexts.

8.1 Addressing Research Questions

This thesis began with a literature review of candidate solutions against misinformation, focussing on game-based inoculation interventions. First, this review revealed a trend of misinformation serious games being digital (Kiili et al., 2024); meanwhile, serious game literature posits that analogue designs may in fact be particularly good platforms for learning (Bayeck, 2020; Fjællingsdal & Klöckner, 2020; Rogerson et al., 2016; Smith & Golding, 2018; Xie et al., 2008). Furthermore, there is little evidence of how the outcomes of game-based interventions (both digital and analogue) change over time (Maertens et al., 2021, 2025). This is important, as if the inoculation effect decays rapidly, this type of intervention may require closer inspection to more robustly protect the individual (Roozenbeek et al., 2024). Recent research by Maertens et al. (2025) has also called for the long-term effectiveness of inoculation to be considered in terms of learning, and inoculation scholarship is yet to comprehensively draw on established models and theories of learning. Finally, the literature review revealed that some scholars and practitioners had been frustrated in the lack of understanding of 'what works' in game-based intervention design (Appel et al., 2025; Roozenbeek et al., 2024; Saleh et al., 2023). This section will address each research question in turn, discussing how the research in this thesis has sought to answer the research questions of this thesis. For the reader's benefit, the three research questions of this thesis are listed below.

RQ1. How effective are analogue game-based interventions against misinformation, and to what extent does this change over time?

RQ2. What kinds of learning are stimulated by analogue game-based inoculation interventions against misinformation?

RQ3. How could game-based inoculation interventions against misinformation be better designed?

The first research question is defined in two parts: the effectiveness of analogue game-based interventions, and how robust this effect is to time. The quantitative findings of the *Fake News* game intervention presented in Chapter 4 showed very little evidence of short term or long term changes in susceptibility to misinformation. Statistical tests demonstrated that across three different within-subject measures the effect of the intervention was not significant. Moreover, differences between a longitudinal post-test (9 weeks post-intervention) and an online control group were also not significant across all three measures. These findings indicate that analogue game-based interventions may not be effective approaches to reducing susceptibility to misinformation. However, the qualitative findings presented in Chapter 5 offer more of an insight into participant 'transformations' from the intervention, finding that players gained knowledge from one another as they completed the task collaboratively leading to outcomes broader than the quantitative measures used. Moreover, participant interviews found that that the collaborative, tactile, and embodied nature of the analogue intervention *Fake News* seemed to make it more memorable, potentially making the effects of the game more durable over time. These findings indicate that analogue game-based inoculations can be effective interventions, and the platform (i.e., analogue) may be particularly effective at reducing the decay of these outcomes. However, scholars must consider that game-based intervention outcomes may be broad and go beyond behavioural 'transformations' in players.

The second research question asks what kinds of learning occur within analogue game-based interventions. The analogue intervention *Fake News* stimulated outcomes that can be interpreted through different learning theories, such as experiential learning, reflective learning, and collaborative learning. Chapter 5 highlights that in-person intervention design may be particularly effective at establishing 'concrete experiences' on which players can reflect and draw new abstract conceptualisations and abstractions. Furthermore, as highlighted in the comments above, session observations repeatedly found that players learned from each other as they worked collaboratively to complete the intervention. Participants were observed sharing relevant misinformation techniques with each other and applying these techniques to real world news publications and politicians. In answering the second

research question this thesis finds that analogue interventions can stimulate a wide range of learning, in particular experiential learning. Chapter 5 argues that learning theories can prove valuable not only in evaluating intervention outcomes, but in designing them too; collaborative learning, for example, can even be purposely and effectively stimulated through collaborative analogue design.

The third research question asks how game-based inoculation interventions may be better designed, and this thesis finds that principally this may be achieved through a) closer interaction to inoculation theory mechanisms, and b) greater co-action with (serious) game design literature. First, despite often being framed as following 'theory-led design', many current game-based interventions do not yet effectively draw on inoculation theory. Although inoculation literature does not yet provide instructive message design protocol for game-based contexts, findings from Chapter 7 demonstrate that conventional message design protocols can still be followed in this context. Importantly, this thesis does not argue that existing interventions are ineffective; instead, Chapter 6 and 7 argue that scholars and practitioners may be able to create more effective interventions by interacting with inoculation theory more comprehensively. Second, Chapter 6 revealed that current game-based inoculation scholarship lacks transparency and discussion surrounding game design and development processes, and argues that greater interaction with (serious) game design and development literature may help guide practitioners in creating these interventions. Serious game design literature may not only help explain why interventions are (in)effective (Roozenbeek et al., 2024) but can demonstrate how to design them more effectively (e.g., Culyba, 2018; Fullerton, 2004). By drawing from this literature, and importantly, maintaining transparency around design and development process, Chapter 6 argues that scholars can facilitate richer comparison between approaches, and guide practitioners to make more effective interventions.

8.2 Key Contributions

Through addressing the three research questions addressed above, this thesis has made novel contributions to game-based interventions against misinformation citing inoculation as their foundational theory. This section presents these findings as being either empirical, theoretical, and methodological. That is, this section not only summarises empirical contributions but also how this thesis has contributed to existing theories and approaches in game-based inoculation research, and the methodological advances of the approach adopted.

Empirical Contributions

This thesis has made several empirical contributions to inoculation literature, particularly around the possible effectiveness and advantages of non-digital platforms

to host game-based interventions. Existing reviews of misinformation serious games (as well as the integrative review in this thesis) have shown that existing interventions are overwhelmingly digital. Serious game literature, however, has suggested that analogue interventions may be more engaging and enjoyable (Rogerson et al., 2016; Taspinar et al., 2016), effective at encouraging collaboration (Berland & Lee, 2011; Xie et al., 2008), and can be effective communication and learning tools (Bayeck, 2020; Fjællingsdal & Klöckner, 2020; Smith & Golding, 2018). This thesis provides contrasting evidence for the efficacy of analogue interventions as learning tools. Quantitative findings suggested that the intervention did not reduce players' perceived persuasiveness, reliability, and personal agreement with 'fake news' articles, representing the second failed replication with this intervention (Roozenbeek & van der Linden, 2018; Wong & Wu, 2023).

However, qualitative findings showed that the analogue collaborative design was enjoyable, resistant to memory fade, and facilitated collaborative learning. Inoculation scholars have recently been investigating inoculation hesitancy, an unwillingness to engage with an intervention (A. Johnson & Madsen, 2024). An intervention can be efficacious but remain ineffective if individuals are unwilling to engage with it (Roozenbeek et al., 2024). Qualitative findings from the study demonstrated that (particularly older) participants enjoyed the analogue collaborative elements of the intervention, and many expressed that this was a leading contributor to them taking part in the study (over financial incentives offered). Participants also expressed that they would be interested in replaying the intervention with family members (particularly children), further emphasising the attractiveness and replay potential of analogue designs in these contexts.

The resistance of the inoculation effect over time has also received attention from scholars (Maertens et al., 2021, 2025), although research is still lacking in this area (Roozenbeek et al., 2024). This is important as the real-world impact of an intervention is severely diminished if its effects decay quickly. Research by Maertens et al. (2021) found that inoculating effect of the digital intervention *Bad News* reduced after 8 weeks. Qualitative interviews in this study demonstrated that participants were able to reliably recall elements of the intervention at least 15 weeks after the intervention had taken place and is an encouraging effect of analogue design. As heralded in the comments above, the study with the analogue intervention *Fake News* also observed that players learned collaboratively from one another during play. This activity of collaboration and problem solving has been observed in conjunction with learning in analogue game design literature before (Berland & Lee, 2011; Peppler et al., 2013; Zagal et al., 2006), but to the authors knowledge has not been explored in game-based inoculation settings. Analogue design may be an effective approach to not only inoculating

individuals through intervention materials but also facilitating players to inoculate one another.

Theoretical Contributions

Chapter 6 and 7 of this thesis have offered theoretical contributions to game-based inoculation literature, particularly on how inoculation theory can be drawn upon in game-based contexts. The integrative narrative review presented in chapter 6 revealed that, although often effective, many existing game-based inoculation interventions seeking to build resistance to misinformation do not comprehensively draw upon inoculation theory literature. The extend to which inoculation theory is drawn upon in many of these game-based contexts is in the theory's pre-emptive delivery, although even then, in corresponding publications authors have highlighted inoculation's efficacy in therapeutic settings (i.e., where the individual has already been exposed to the attack, but see Section 2.2 on other definitions of this). This thesis has demonstrated, however, that it is possible to follow established inoculation message design protocol in game-based contexts. Refutational pre-emption, a message component in conventional interventions, can form a central mechanic in a game-based setting as players actively interact with 'weakened dose' attack messages and their refutations. Such an intervention may be described as being both issue-based and technique-based in design; the intervention contains pre-emptive refutational messages (a feature of conventional issue-based interventions), yet refutations inform the individual of the techniques used in the message.

This thesis has highlighted further theoretical shortcomings in current game-based intervention design. Game-based intervention design is a highly complex process (Culyba, 2018) and is highly interdisciplinary in nature, and draws from not only communication and learning theories, but requires technical and design skills as well. Yet, this thesis has identified that game design and development literature – an important process that plays a central role in influencing the final intervention (Fullerton, 2004; Salen & Zimmerman, 2004; Schell, 2019) – has not yet been drawn comprehensively upon in the majority of cases. A game-based intervention may contain pre-emptive refutational messages, but if the intervention does not engage or motivate the player, these may not have the intended effect. This is particularly true of analogue collaborative games, which must also motivate social and collaborative actions between players (Zagal et al., 2006). Serious game design frameworks such as The Transformational Framework (Culyba, 2018) and game development processes such as 'playcentric' design may, as demonstrated in Chapter 7 of this thesis, provide instructional support to practitioners and scholars. Chapter 6 of this thesis found that through an iterative test-evaluate-revise processes an intervention's design may be

improved and refined. Engagement with this literature may not only help create practitioners design more effective interventions.

Following calls for a more learning-oriented consideration of inoculation theory (Maertens et al., 2025), this thesis sought to draw on existing learning theory literature. As heralded in the comments above, qualitative findings found that the experiences of players shared many similarities with existing learning theories and frameworks (e.g., experiential learning, reflective learning, and collaborative learning). This thesis finds that learning theories can also be drawn upon when designing interventions. For example, collaborative learning literature played a role in the design of the new intervention presented in Chapter 7, *The disPHISHinformation Game*. Learning theories may also be particularly useful to practitioners designing technique-based inoculation interventions, where a message design protocol has yet to be established.

Methodological Contributions

A major methodological contribution of this thesis to inoculation theory literature was the use of qualitative methods. Following an inductive pragmatist research paradigm (see Chapter 3), I decided to break with recent convention[18] by including qualitative session observations and semi-structured interviews. These methods have played a major role in not only helping to understand the (often varied) outcomes of the intervention but deciphering why the intervention elicited these player transformations. Moreover, adopting this mixed-methods approach allowed qualitative findings to contextualise quantitative results. Recently scholars have expressed frustration in not understanding why interventions have failed (Harjani et al., 2023), or even why they have succeeded (Roozenbeek et al., 2024). Following a positivist research philosophy has meant that interventions have been treated as 'black box' systems. That is, quantitative methods have required researchers to consider interventions only in terms of their output. This thesis, however, has demonstrated that (particularly analogue) game-based interventions involve internal (often social) mechanisms that may be highly consequential in intervention outcomes. Observing these processes may not only contextualise quantitative results but help practitioners improve intervention design.

Finally, a key finding from the first study of this thesis was that the outcomes from the analogue intervention were broad, going beyond the quantitative measures used to measure intervention efficacy. While the quantitative measures deployed measured primarily behavioural outcomes from the intervention, qualitative methods

[18] Although inoculation research has in the past used thought-listing or think aloud processes in measuring counterarguing (see Ivanov et al., 2013).

revealed that the intervention also improved knowledge of misinformation techniques and, as an experience, changed players' perspectives on misinformation techniques on which they were able to reflect. In fact, serious game design literature suggests that outcomes from serious games can also influence beliefs, relationships, sense of identity, and even physical health (Culyba, 2018). Considering the types of transformations stimulated by an intervention can not only influence the intervention's design, but also how the outcomes of the intervention are measured. Following pragmatist methodologies to measure, for example, changes in perspectives, affective outcomes, and altered beliefs, may be essential to not only understanding intervention effectiveness, but *how* the intervention has stimulated these outcomes.

8.3 Future Work

The findings from this thesis stress the importance of taking intervention design considerations and development processes seriously (in the context of game-based interventions against misinformation) and highlights the impact that this may have on intervention outcomes. This, as well as the novel methods used in this thesis, are central to the contributions of this thesis, and carve out a novel research agenda around game-based inoculation. The implications of the findings and contributions presented in this chapter are important for existing practitioners designing game-based inoculation interventions and may help guide future research in this area. Drawing from the contributions of this thesis, this section outlines an agenda for research and practice which aims to inform the creation of more effective and attractive game-based interventions against misinformation and further topics as well.

First, the adoption of an inductive pragmatist research philosophy (and subsequent use of mixed methods) afforded a far more comprehensive understanding of intervention outcomes with the intervention *Fake News* than through quantitative measures alone. Scholars have called for more research understanding *why* interventions are (in)effective, yet qualitative methods remain broadly unexplored (Roozenbeek et al., 2024). Scholars should consider integrating qualitative elements when designing game-based intervention studies, as these methods may not only help measure changes to behaviour but wider player transformations in knowledge, skill, beliefs, and affect (Culyba, 2018). Recent replications and reanalyses of game-based interventions have raised questions on efficacy (M. E. Graham et al., 2023; Modirrousta-Galian & Higham, 2023); this thesis has shown that qualitative methods may be particularly useful as tools to investigate these claims. Participant interviews can be powerful tools to measure intervention outcomes contextualised within players' personal experiences and engagements with misinformation. Moreover, qualitative methods such as session observations allow researchers to begin to consider the processes occurring during the intervention, particularly with analogue intervention

design. Beyond the scope of the research carried out in this thesis, future research could explore additional qualitative methods such as focus groups for creating additional participant dialogue and interchange, or methods for deriving participant elicited data (such as diaries, photos or videos). Scholars should use these tools to further explore the efficacy of analogue interventions, an under-explored but promising platform for learning and inoculation.

Second, this thesis has demonstrated a novel approach to drawing on inoculation in game-based contexts that more comprehensively follows the theory's established message design protocols (e.g., Ivanov, 2017). Chapter 6 demonstrated that many existing game-based interventions draw on inoculation theory superficially, and without established literature on game-based message design the interpretation of inoculation mechanisms in these game-based contexts has been largely down to the researcher or practitioner's own discretion. This thesis has demonstrated that this need not be the case, and that even in game-based contexts conventional inoculation message design protocols can still be drawn upon. This also opens an exciting avenue of future research for inoculation; rather than game-based interventions being auxiliary to inoculation, they may now be more sincerely embedded in the literature, contributing to our understanding of inoculation itself. For example, research shows that game-based platforms may be uniquely poised to experiment with eliciting affect in players (e.g., through narrative transportation, see Frome, 2007), an important moderator in inoculation and a variable linked to the threat mechanism. Having established that game-based interventions can have a deeper foundational grounding in inoculation theory through the use of conventional components, practitioners could explore creative approaches to delivering these components through game mechanics. Deliberate and active experimentation with the different mechanisms and components of inoculation could be integrated into game design processes in order to assess effectiveness and explore alternative approaches. Moreover, practitioners may now be able to experiment with more validity how inoculation can be used in collaboration with other communication theories (e.g., transportation theory, see Grace & Liang, 2023), or even learning theories (e.g., experiential or collaborative learning), as demonstrated in this thesis. As highlighted by Maertens et al. (2025), coaction with memory and learning literature may support stronger intervention outcomes that last for longer.

Finally, this thesis has demonstrated the wide scope of scholarship possible in the design and development of game-based interventions. For many existing game-based interventions, inoculation has been framed playing a central role in their design (e.g., Appel et al., 2025; Basol et al., 2021; Roozenbeek & van der Linden, 2018). Game design literature tells us that the design and development processes themselves are

central in shaping the interventions (Culyba, 2018; Fullerton, 2004; Salen & Zimmerman, 2004), but this literature has not yet been drawn upon in existing game-based inoculation scholarship (see Chapter 6). Yet, many existing interventions have been designed and developed by (or in collaboration with) established studios such as Gusmanson[19], TILT[20], and DROG[21], who will have in all likelihood followed iterative development processes (State of Agile, 2023). Scholars should adopt a more transparent approach with regards to design and development processes when presenting new interventions, as these processes may have contributed equally to the intervention's outcomes as inoculation theory. Promising work by Cook et al. (2024) has discussed how their adoption of co-design development approaches has influenced intervention design, but this dialogue is uncommon. Game and intervention design itself should become an important aspect of exploration in this literature. To do this, studies must more explicitly experiment with and compare different game design decisions (e.g., analogue vs digital, collaborative vs individual) and game design processes.

More transparency and discussion around design and development processes can begin dialogues on the advantages of certain approaches and may begin to build game-based inoculation intervention design and development protocols. Compared to conventional inoculation messages game-based interventions are highly complex, and scholars must begin to draw on (serious) game design and development literature to facilitate this process. Serious games are highly promising platforms for hosting inoculation interventions, and research has demonstrated that they can be effective at improving player resilience to misinformation. Some scholars have expressed frustration at not understanding 'what works' in the design of these interventions, and this has limited improvement in intervention effectiveness (Appel et al., 2025), or even led to null results (Harjani et al., 2023; Rędzio et al., 2023). Informed from the findings of this thesis, this section has sought to outline a research agenda that may advance game-based inoculation scholarship and ultimately lead to the development of more effective interventions with positive real-world impacts.

[19] https://www.gusmanson.nl/
[20] https://www.tiltstudio.co/
[21] https://drog.group/

9. Bibliography

Aarts, A. A., Anderson, J. E., Anderson, C. J., Attridge, P. R., Attwood, A., Axt, J., Babel, M., Bahník, Š., Baranski, E., Barnett-Cowan, M., Bartmess, E., Beer, J., Bell, R., Bentley, H., Beyan, L., Binion, G., Borsboom, D., Bosch, A., Bosco, F. A., … Zuni, K. (2015). Estimating the reproducibility of psychological science. *Science*, 349(6251). https://doi.org/10.1126/science.aac4716

Abels, C. M., Huttunen, K. J. A., Hertwig, R., & Lewandowsky, S. (2024). Dodging the autocratic bullet: enlisting behavioural science to arrest democratic backsliding. *Behavioural Public Policy*, 1–28. https://doi.org/10.1017/bpp.2024.43

Abt, C. C. (1970). *Serious games*. Viking Press.

Adams, A. D. (2023). Werewolf on campus: A case study in inoculation theory and gamified one-shot library instruction. *Public Services Quarterly*, 19(4), 255–272. https://doi.org/10.1080/15228959.2023.2252322

Agarwal, N., Bandeli, K., Bertolin, G., Biteniece, N., & Sedova, K. (2017). *Digital hydra: Security implications of false information online*. Riga: NATO Strategic Communications Centre of Excellence. Retrieved from https://stratcomcoe.org/publications/digital-hydra-security-implications-of-false-information-online/205

Alabdan, R. (2020). Phishing attacks survey: Types, vectors, and technical approaches. *Future Internet*, 12(10), 1–39. https://doi.org/10.3390/fi12100168

Aladawy, D., Beckers, K., & Pape, S. (2018). PERSUADED: Fighting social engineering attacks with a serious game. In S. Furnell, H. Mouratidis, & G. Pernul (Eds.). *Trust, privacy and security in digital business* (pp. 103–118). Springer International Publishing.

Alaphilippe, A., Gizikis, A., Hanot, C., & Bontcheva, K. (2019). *Automated tackling of disinformation*. European Science-Media Hub. Retrieved from https://www.disinfo.eu/publications/automated-tackling-of-disinformation/

Aleroud, A., & Zhou, L. (2017). Phishing environments, techniques, and countermeasures: A survey. *Computers & Security*, 68, 160-196. https://doi.org/10.1016/j.cose.2017.04.006

Aleven, V., Myers, E., Easterday, M., & Ogan, A. (2010). Toward a framework for the analysis and design of educational games. *IEEE International Conference on Digital Game and Intelligent Toy Enhanced Learning*, 69–76. https://doi.org/10.1109/DIGITEL.2010.55

Alfano, M., Fard, A. E., Carter, J. A., Clutton, P., & Klein, C. (2021). Technologically scaffolded atypical cognition: the case of YouTube's recommender system. *Synthese*, 199(1), 835–858. https://doi.org/10.1007/s11229-020-02724-x

Alkaissi, H., & McFarlane, S. I. (2023). Artificial hallucinations in ChatGPT: Implications in scientific writing. *Cureus*, 15(2), e35179. https://doi.org/10.7759/cureus.35179

Allen, J., Arechar, A. A., Pennycook, G., & Rand, D. G. (2021). Scaling up fact-checking using the wisdom of crowds. *Science Advances*, 7(36), eabf4393. https://doi.org/10.1126/sciadv.abf4393

Allen, J., Howland, B., Mobius, M., Rothschild, D., & Watts, D. J. (2020). Evaluating the fake news problem at the scale of the information ecosystem. *Science Advances*, 6(14), eaay3539. https://doi.org/10.1126/sciadv.aay3539

Altay, S. (2023). Who is concerned about misinformation and why? Evidence from 46 countries between 2018 and 2023. *Information, Communication & Society*, 1-17. https://doi.org/10.31234/osf.io/fz6wu

Altay, S., Berriche, M., & Acerbi, A. (2023). Misinformation on Misinformation: Conceptual and Methodological Challenges. *Social Media + Society*, 9(1). https://doi.org/10.1177/20563051221150412

Alvarez, J., Irrmann, O., Djaouti, D., Taly, A., Rampnoux, O., & Sauvé, L. (2019). Design Games and Game Design: Relations Between Design, Codesign and Serious Games in Adult Education. In *From UXD to LivXD* (pp. 229–253). Wiley. https://doi.org/10.1002/9781119612254.ch11

Anania, E. C., Keebler, J. R., Anglin, K. M., & Kring, J. P. (2016). Using the cooperative board game pandemic to study teamwork. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 60(1), 1770–1774. https://doi.org/10.1177/1541931213601405

Anderson, C. W. (2021). Fake News is Not a Virus: On Platforms and Their Effects. *Communication Theory*, 31(1), 42–61. https://doi.org/10.1093/ct/qtaa008

Anderson, J. R. (1985). *Cognitive psychology and its implications*. W.H. Freeman & Co.

Annetta, L. A. (2010). The "I's" have it: A framework for serious educational game design. *Review of General Psychology*, 14(2), 105–112. https://doi.org/10.1037/a0018985

Anselm, S., & Corbin, J. (1990). *Basics of qualitative research: Techniques and procedures for developing grounded theory*. Sage Publications.

Appel, R. E., Roozenbeek, J., Rayburn-Reeves, R., Basol, M., Corbin, J., Compton, J., & van der Linden, S. (2025). Psychological inoculation improves resilience to and reduces willingness to share vaccine misinformation. *Scientific Reports*, 15(1), 2045–2322. https://doi.org/10.1038/s41598-025-09462-5

Ávila-Pesántez, D., Rivera, L. A., & Alban, M. S. (2017). Approaches for Serious game design: A systematic literature review. *Computers In Education Journal*, 8(3). Retrieved from https://research.ebsco.com/linkprocessor/plink?id=468547b3-2b88-3688-89ea-5c0f556b1b0c

Avison, D. E., Lau, F., Myers, M. D., & Nielsen, P. A. (1999). Action research. *Communications of the ACM*, 42(1), 94–97. https://doi.org/10.1145/291469.291479

Axelsson, C. A. W., Guath, M., & Nygren, T. (2021). Learning how to separate fake from real news: Scalable digital tutorials promoting students' civic online reasoning. *Future Internet*, 13(3), 1–18. https://doi.org/10.3390/fi13030060

Axelsson, C. A. W., Nygren, T., Roozenbeek, J., & van der Linden, S. (2024). Bad News in the civics classroom: How serious gameplay fosters teenagers' ability to discern misinformation techniques. *Journal of Research on Technology in Education*, 0(0), 1–27. https://doi.org/10.1080/15391523.2024.2338451

Badshah, N. (2025). Hackers reportedly steal pictures of 8,000 children from Kido nursery chain. *The Guardian*. Retrieved from https://www.theguardian.com/technology/2025/sep/25/cybercriminals-steal-pictures-and-details-of-8000-children-from-nursery-chain

Bago, B., Rand, D. G., & Pennycook, G. (2020). Fake news, fast and slow: Deliberation reduces belief in false (but not true) news headlines. *Journal of Experimental Psychology: General*, 149(8), 1608–1613. https://doi.org/10.1037/xge0000729

Bail, C. A., Argyle, L. P., Brown, T. W., Bumpus, J. P., Chen, H., & Hunzaker, M. B. F. (2018). Exposure to opposing views on social media can increase political polarization. *Proceedings of the National Academy of Sciences of the United States of America*, 115(37). https://doi.org/10.1073/pnas.1804840115

Bain, J. D., Ballantyne, R., Packer, J., & Mills, C. (1999). Using journal writing to enhance student teachers' reflectivity during field experience placements. *Teachers and Teaching*, *5*(1), 51–73. https://doi.org/10.1080/1354060990050104

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

Banas, J. A., & Richards, A. S. (2017). Apprehension or motivation to defend attitudes? Exploring the underlying threat mechanism in inoculation-induced resistance to persuasion. *Communication Monographs*, 84(2), 164–178. https://doi.org/10.1080/03637751.2017.1307999

Bandura, A. (1977). *Social learning theory*. Prentice-Hall.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Prentice Hall.

Bandura, A., & Cervone, D. (1983). Self-evaluative and self-efficacy mechanisms governing the motivational effects of goal systems. *Journal of Personality and Social Psychology*, *45*(*5*), 1017–1028. https://doi.org/10.1037/0022-3514.45.5.1017

Barab, S. A., Scott, B., Siyahhan, S., Goldstone, R., Ingram-Goble, A., Zuiker, S. J., & Warren, S. (2009). Transformational play as a curricular scaffold: Using videogames to support science education. *Journal of Science Education and Technology*, 18(4), 305–320. https://doi.org/10.1007/s10956-009-9171-*5*

Bartlett, F. (1932). *Remembering: A study in experimental and social psychology*. Cambridge University Press.

Barzilai, S., & Chinn, C. A. (2018). On the goals of epistemic education: Promoting apt epistemic performance. *Journal of the Learning Sciences*, 27(3), 353–389. https://doi.org/10.1080/10508406.2017.1392968

Basol, M. (2022). Essays on resistance against persuasion: Building, strengthening, and spreading attitudinal resistance through inoculation theory [Doctoral dissertation, University of Cambridge]. Apollo Cambridge Repository. https://doi.org/10.17863/CAM.89417

Basol, M., Roozenbeek, J., & Van Der Linden, S. (2020). Good news about bad news: Gamified inoculation boosts confidence and cognitive immunity against fake news. *Journal of Cognition*, 3(1). https://doi.org/10.5334/joc.91

Basol, M., Roozenbeek, J., Berriche, M., Uenal, F., McClanahan, W. P., & Linden, S. van der. (2021). Towards psychological herd immunity: Cross-cultural evidence for two prebunking interventions against COVID-19 misinformation. *Big Data and Society*, 8(1). https://doi.org/10.1177/20539517211013868

Bayeck, R. Y. (2020). Examining board gameplay and learning: A multidisciplinary review of recent research. *Simulation and Gaming*, 51(4), 411–431. https://doi.org/10.1177/1046878119901286

BBC. (2020). Ofcom: Covid-19 5G theories are "most common" misinformation. Retrieved from https://www.bbc.co.uk/news/technology-52370616

Becker, M. H., Williams, M. J., & Hassan, A. (2025). Examining the impacts of a gamified media literacy intervention in Indonesia. *Journal for Deradicalization*, 43, 140–175. Retrieved from https://jd.journals.publicknowledgeproject.org/jd/index.php/jd/article/view/1055

Bednarek, A. T., Miyamoto, B., Corbett, K., Hudson, C., Scarrow, G., Brass, M., DuMont, K., Holmes, B., Supplee, L., McLean, R., Bhatia, J., Chappell, S., Vélez, M. G., & Kolavalli, C. (2025). How and why funders support engaged research. *Proceedings of the National Academy of Sciences of the United States of America*, 122(1), 1–9. https://doi.org/10.1073/pnas.2400931121

Bengston, D. N., Westphal, L. M., Dockry, M. J., & Crabtree, J. (2022). A "serious game" to explore alternative forestry futures. *Journal of Forestry*, 120(2), 222–226. https://doi.org/10.1093/jofore/fvab059

Berland, M., & Lee, V. R. (2011). Collaborative strategic board games as a site for distributed computational thinking. *International Journal of Game-Based Learning*, 1(2), 65–81. https://doi.org/10.4018/ijgbl.2011040105

Bessi, A., & Ferrara, E. (2016). Social bots distort the 2016 US presidential election online discussion. Retrieved from https://ssrn.com/abstract=2982233

Blair, R. A., Gottlieb, J., Nyhan, B., Paler, L., Argote, P., & Stainfield, C. J. (2024). Interventions to counter misinformation: Lessons from the Global North and applications to the Global South. *Current Opinion in Psychology*, *55*, 101732. https://doi.org/10.1016/j.copsyc.2023.101732

Blinder, S., & Allen, W. L. (2016). *UK public opinion toward immigration: Overall attitudes and level of concern*. The Migration Observatory (University of Oxford). Retrieved from www.migrationobservatory.ox.ac.uk

Bloomfield, S. F., Stanwell-Smith, R., Crevel, R. W. R., & Pickup, J. (2006). Too clean, or not too clean: The hygiene hypothesis and home hygiene. *Clinical and Experimental Allergy*. 36(4), 402–425. https://doi.org/10.1111/j.1365-2222.2006.02463.x

Bolsen, T., Druckman, J. N., & Lomax, F. (2014). The influence of partisan motivated reasoning on public opinion. *Political Behaviour*, 36(2), 235–262. https://doi.org/10.1007/s11109-013-9238-0

Bonnett, A., & Hopkins, P. (2025). Understanding the 2024 UK riots. *The Geographical Journal*, e70047. https://doi.org/10.1111/geoj.70047

Booker, K. L., & MItchell, A. W. (2021). From boring to board game: The effect of a serious game on key learning outcomes. *Journal of Occupational Therapy* Education, *5*(4). https://doi.org/10.26681/jote.2021.050407

Boud, D. (1990). Assessment and the promotion of academic values. *Studies in Higher Education*, 15(1), 101–111. https://doi.org/10.1080/03075079012331377621

Boud, D., Keogh, R., & Walker, D. (Eds.). (1985). *Reflection: turning experience into learning*. Routledge

Boyd, E. M., & Fales, A. W. (1983). Reflective learning: Key to learning from experience. *Journal of Humanistic Psychology*, 23, 99-117. https://doi.org/10.1177/00221678832320

Boyle, E. A., Hainey, T., Connolly, T. M., Gray, G., Earp, J., Ott, M., Lim, T., Ninaus, M., Ribeiro, C., & Pereira, J. (2016). An update to the systematic literature review of empirical evidence of the impacts and outcomes of computer games and serious games. *Computers and Education*, 94, 178–192. https://doi.org/10.1016/j.compedu.2015.11.003

Boyle, E., Connolly, T. M., & Hainey, T. (2011). The role of psychology in understanding the impact of computer games. *Entertainment Computing*, 2(2), 69–74. https://doi.org/10.1016/j.entcom.2010.12.002

Brady, W. J., Crockett, M. J., & Van Bavel, J. J. (2020). The MAD model of moral contagion: The role of motivation, attention, and design in the spread of moralized content online. *Perspectives on Psychological Science*, 15(4), 978–1010. https://doi.org/10.1177/1745691620917336

Brady, W. J., Gantman, A. P., & Van Bavel, J. J. (2020). Attentional capture helps explain why moral and emotional content go viral. *Journal of Experimental Psychology: General*, 149(4), 746–756. https://doi.org/10.1037/xge0000673

Brady, W. J., Wills, J. A., Jost, J. T., Tucker, J. A., Van Bavel, J. J., & Fiske, S. T. (2017). Emotion shapes the diffusion of moralized content in social networks. *Proceedings of the National Academy of Sciences of the United States of America*, 114(28), 7313–7318. https://doi.org/10.1073/pnas.1618923114

Brandt, M. J., IJzerman, H., Dijksterhuis, A., Farach, F. J., Geller, J., Giner-Sorolla, R., Grange, J. A., Perugini, M., Spies, J. R., & van 't Veer, A. (2014). The Replication Recipe: What makes for a convincing replication? *Journal of Experimental Social Psychology*, 50(1), 217–224. https://doi.org/10.1016/j.jesp.2013.10.005

Brandtzaeg, P. B., Følstad, A., & Chaparro Domínguez, M. Á. (2018). How journalists and social media users perceive online fact-checking and verification services. *Journalism Practice*, 12(9), 1109–1129. https://doi.org/10.1080/17512786.2017.1363657

Bronstein, M. V., Pennycook, G., Bear, A., Rand, D. G., & Cannon, T. D. (2019). Belief in fake news is associated with delusionality, dogmatism, religious fundamentalism, and reduced analytic thinking. *Journal of Applied Research in Memory and Cognition*, 8(1), 108–117. https://doi.org/10.1016/j.jarmac.2018.09.005

Bryman, A. (2012). *Social research methods* (4th ed.). Oxford University Press.

Burnell, P., & PA Media. (2024). No charge over spreading of Southport misinformation. *BBC News*. Retrieved from https://www.bbc.co.uk/news/articles/crl8nwx6ynzo

Burgoon, M., Miller, M. D., Cohen, M., & Montgomery, C. L. (1978). An empirical test of a model of resistance to persuasion. *Human Communication Research*, 5(1), 27–39. https://doi.org/10.1111/j.1468-2958.1978.tb00620.x

Burkill, S., Copas, A., Couper, M. P., Clifton, S., Prah, P., Datta, J., Conrad, F., Wellings, K., Johnson, A. M., & Erens, B. (2016). Using the web to collect data on sensitive behaviours: A study looking at mode effects on the British national survey of sexual attitudes and lifestyles. *PLoS ONE*, 11(2). https://doi.org/10.1371/journal.pone.0147983

Cameron, B., & Dwyer, F. (2005). The effect of online gaming, cognition and feedback type in facilitating delayed achievement of different learning objectives. *Journal of Interactive Learning Research*, 16(3), 243–258. Retrieved from https://www.learntechlib.org/p/5896

Capewell, G., Maertens, R., Remshard, M., van der Linden, S., Compton, J., Lewandowsky, S., & Roozenbeek, J. (2024). Misinformation interventions decay rapidly without an immediate posttest. *Journal of Applied Social Psychology*, *54*(8), 441–454. https://doi.org/10.1111/jasp.13049

Carlson, M. (2020). Fake news as an informational moral panic: the symbolic deviancy of social media during the 2016 US presidential election. *Information Communication and Society*, 23(3), 374–388. https://doi.org/10.1080/1369118X.2018.1505934

Center for Countering Digital Hate. (2024). X ran ads on five accounts pushing lies and hate during UK riots. *Center for Countering Digital Hate*. Retrieved from https://counterhate.com/research/x-ran-ads-on-five-accounts-pushing-lies-and-hate-during-uk-riots/

Celadin, T., Capraro, V., Pennycook, G., & Rand, D. G. (2023). Displaying news source trustworthiness ratings reduces sharing intentions for false news posts. *Journal of Online Trust and Safety*, 1(5), 1–20. https://doi.org/10.54501/jots.v1i5.100

Cellan-Jones, R. (2018). Fake news: Can teenagers spot it? *BBC News*. Retrieved from https://www.bbc.co.uk/news/technology-46206675

Censorious governments are abusing "fake news" laws. (2021). *The Economist*. Retrieved from https://www.economist.com/international/2021/02/13/censorious-governments-are-abusing-fake-news-laws

Chan, M. pui S., & Albarracín, D. (2023). A meta-analysis of correction effects in science-relevant misinformation. *Nature Human Behaviour*, 7(9), 1514–1525. https://doi.org/10.1038/s41562-023-01623-8

Cheshire, T., & Doak, S. (2024). Southport attack misinformation fuels far-right discourse on social media. *Sky News*. Retrieved from https://news.sky.com/story/southport-attack-misinformation-fuels-far-right-discourse-on-social-media-13188274

Chiew, K. L., Yong, K. S. C., & Tan, C. L. (2018). A survey of phishing attacks: Their types, vectors and technical approaches. *Expert Systems with Applications*, 106, 1–20. https://doi.org/10.1016/j.eswa.2018.03.050

Compton, J. (2009). Threat explication: What we know and don't yet know about a key component of inoculation theory. *STAM Journal*, 39, 1–18.

Compton, J. (2012). Inoculation theory. In J. P. Dillard & L. Shen (Eds.), *The SAGE handbook of persuasion: Developments in theory and practice* (2nd ed., pp. 220–236). Sage.

Compton, J. (2018). Inoculating against / with political humor. In J. C. Baumgartner & A. B. Becker (Eds.), *Political humor in a changing media landscape: A new generation of research* (pp. 95 – 113). Lexington Books.

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. (2021). Threat and / in Inoculation Theory. *International Journal of Communication*, 15, 4294–4306. Retrieved from https://ijoc.org/index.php/ijoc/article/view/17634

Compton, J. (2024). Inoculation Theory. *Review of Communication*, 25(1), 1–13. https://doi.org/10.1080/15358593.2024.2370373

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. (2005). Inoculation theory of resistance to influence at maturity: Recent progress in theory development and application and suggestions for future research. *Annals of the International Communication Association*, 29(1), 97–146. https://doi.org/10.1080/23808985.2005.11679045

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

Compton, J., Ivanov, B., & Hester, E. (2022). Inoculation theory and affect. *International Journal of Communication*, 16, 3470–3483. Retrieved from https://ijoc.org/index.php/ijoc/article/view/19094

Compton, J., van der Linden, S., Cook, J., & Basol, M. (2021). Inoculation theory in the post-truth era: Extant findings and new frontiers for contested science, misinformation, and conspiracy theories. *Social and Personality Psychology Compass*, 15(6), e12602. https://doi.org/10.1111/spc3.12602

Connolly, T. M., Boyle, E. A., MacArthur, E., Hainey, T., & Boyle, J. M. (2012). A systematic literature review of empirical evidence on computer games and serious games. *Computers and Education*, *5*9(2), 661–686. https://doi.org/10.1016/j.compedu.2012.03.004

Conover, M., Ratkiewicz, J., Francisco, M., Goncalves, B., Menczer, F., & Flammini, A. (2021). Political Polarization on Twitter. *Proceedings of the International AAAI Conference on Web and Social Media*, *5*(1), 89–96. https://doi.org/10.1609/icwsm.v5i1.14126

Cook, J. (2021). Cranky Uncle: A game building resilience against climate misinformation. *Plus Lucis*, 3, 13–16. Retrieved from https://crankyuncle.com/wp-content/uploads/2021/10/Cook_2021_Cranky_Uncle.pdf

Cook, J., Ecker, U. K. H., Trecek-King, M., Schade, G., Jeffers-Tracy, K., Fessmann, J., Kim, S. C., Kinkead, D., Orr, M., Vraga, E., Roberts, K., & McDowell, J. (2023). The cranky uncle game—combining humor and gamification to build student resilience against climate misinformation. *Environmental Education Research*, 29(4), 607–623. https://doi.org/10.1080/13504622.2022.2085671

Cook, J., Lepage, C., Hopkins, K. L., Cook, W., Kolog, E. A., Thomson, A., Iddrisu, I., & Burnette, S. (2024). Co-designing and pilot testing a digital game to improve vaccine attitudes and misinformation resistance in Ghana. *Human Vaccines & Immunotherapeutics*, 20(1), 2407204. https://doi.org/10.1080/21645515.2024.2407204

Cook, J., Lewandowsky, S., & Ecker, U. K. H. (2017). Neutralizing misinformation through inoculation: Exposing misleading argumentation techniques reduces their influence. *PLoS ONE*, 12(*5*), 1–21. https://doi.org/10.1371/journal.pone.0175799

Cook, J., Nyirazinyoye, L., Mugisha, M., Uwamahoro, A., Claude, R. J., Lepage, C., Thomson, A., Hopkins, K. L., & Cook, W. (2025). Pilot testing a French and Kinyarwanda game to improve vaccine attitudes and misinformation resilience in Rwanda. *Medical Research Archives*, 13(3). https://doi.org/10.18103/mra.v3i3.6421

Coutinho, S., Bollen, A., Weil, C., Sheerin, C., Silvera, D., Donaldson, S., & Rosborough, J. (2023). *Cyber security skills in the UK labour market 2023: Findings report*. UK Department for Science, Innovation, and Technology. Retrieved from https://www.gov.uk/government/publications/cyber-security-skills-in-the-uk-labour-market-2023

Cowley, B., Heikura, T., & Ravaja, N. (2013). Learning loops - interactions between guided reflection and experience-based learning in a serious game activity. *Journal of Computer Assisted Learning*, 29(4), 348–370. https://doi.org/10.1111/jcal.12013

Cross, L., Piovesan, A., Sousa, M., Wright, P., & Atherton, G. (2023). Your move: An open access dataset of over 1500 board gamer's demographics, preferences and motivations. *Simulation and Gaming*, *54*(*5*), *554–575*. https://doi.org/10.1177/10468781231189493

Culyba, S. (2018). *The Transformational Framework: A process tool for the development of Transformational games*. Carnegie Mellon University.

Cunliffe, A. L., & Scaratti, G. (2017). Embedding impact in engaged research: Developing socially useful knowledge through dialogical sensemaking. *British Journal of Management*, 28(1), 29–44. https://doi.org/10.1111/1467-8551.12204

Curry, J., & Drage, N. (2020). *The handbook of cyber wargames: Wargaming the 21st century*. Independently Published.

Cutumisu, M., Patel, S. D., Brown, M. R. G., Fray, C., von Hauff, P., Jeffery, T., & Schmölzer, G. M. (2019). RETAIN: A board game that improves neonatal resuscitation knowledge retention. *Frontiers in Pediatrics*, 7. https://doi.org/10.3389/fped.2019.00013

Damaševičius, R., Maskeliūnas, R., & Blažauskas, T. (2023). Serious games and gamification in healthcare: A meta-review. *Information*, 14(2), 105. https://doi.org/10.3390/info14020105

Davies, B., Turner, M., & Udell, J. (2024). It helps to be funny or compassionate: An exploration of user experiences and evaluation of social media micro-intervention designs for protecting body image. *Computers in Human Behavior*, 150, 107999. https://doi.org/10.1016/j.chb.2023.107999

DeJong, S. (2023). Playing with fake news: State of fake news video games. *International Journal of Games and Social Impact*, 1(1), 94–111. https://doi.org/10.24140/ijgsi.v1.n1.05

Del Vicario, M., Vivaldo, G., Bessi, A., Zollo, F., Scala, A., Caldarelli, G., & Quattrociocchi, W. (2016). Echo chambers: Emotional contagion and group polarization on Facebook. *Scientific Reports*, 6, 37825. https://doi.org/10.1038/srep37825

DellaVigna, S., & Linos, E. (2022). RCTs to scale: Comprehensive evidence from two nudge units. *Econometrica*, 90(1), 81–116. https://doi.org/10.3982/ecta18709

Denning, T., Lerner, A., Shostack, A., & Kohno, T. (2013). Control-Alt-Hack: The design and evaluation of a card game for computer security awareness and education. *Proceedings of the ACM Conference on Computer and Communications Security*, 915–928. https://doi.org/10.1145/2508859.2516753

Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2015). From game design elements to gamefulness: Defining "gamification." *Library Technology Reports*, *51*(2), *5*–9. https://doi.org/10.1145/2181037.2181040

Deutsch, M. (1949). A theory of co-operation and competition. *Human Relations*, 2(2), 129–152. https://doi.org/10.1177/001872674900200204

Devasia, N., & Lee, J. H. (2024). The role of narrative in misinformation games. *Harvard Kennedy School Misinformation Review*, *5*(5). https://doi.org/10.37016/mr-2020-158

Dewey, J. (1933). *How we think: A restatement of the relation of reflective thinking to the educative process* (2nd ed.). D.C. Heath & Co Publishers.

Dewey, J. (1938). *Experience and education*. Kappa Delta Pi.

Dhillon, G., & Backhouse, J. (2001). Current directions in IS security research: towards socio-organizational perspectives. *Information Systems Journal*, 11(2), 85–178. https://doi.org/10.1046/j.1365-2575.2001.00099.x

Dillard, J. P. (1998). The role of affect in communication, biology, and social relationships. In P. A. Anderson, L. K. Guerrero (Eds.). *Handbook of communication and emotion: Research, theory, applications, and contexts* (pp. 17–32). Academic Press.

Dillenbourg, P. (1999). What do you mean by collaborative learning? In P. Dillenbourg (Ed.). *Collaborative-learning: Cognitive and computational approaches* (2nd ed., pp. 1–19). Emerald Publishing Limited.

Doroshenko, L., & Lukito, J. (2021). Trollfare: Russia's disinformation campaign during military conflict in Ukraine. *International Journal of Communication*, 15, 4662–4689. Retrieved from https://ijoc.org/index.php/ijoc/article/view/16895.

Eagly, A. H., & Chaiken, S. (1993). *The psychology of attitudes*. Harcourt Brace Jovanovich College Publishers.

Earn money from more of your content: Introducing Facebook's new monetization program. (2024). *Meta Platforms Inc*. Retrieved from https://creators.facebook.com/introducing-facebook-content-monetization

Echeverría, A., García-Campo, C., Nussbaum, M., Gil, F., Villalta, M., Améstica, M., & Echeverría, S. (2011). A framework for the design and integration of collaborative

classroom games. *Computers and Education*, *57*(1), 1127–1136. https://doi.org/10.1016/j.compedu.2010.12.010

Ecker, U. K. H., Lewandowsky, S., & Apai, J. (2011). Terrorists brought down the plane!-No, actually it was a technical fault: Processing corrections of emotive information. *Quarterly Journal of Experimental Psychology*, 64(2), 283–310. https://doi.org/10.1080/17470218.2010.497927

Ecker, U. K. H., Lewandowsky, S., Cook, J., Schmid, P., Fazio, L. K., Brashier, N., Kendeou, P., Vraga, E. K., & Amazeen, M. A. (2022). The psychological drivers of misinformation belief and its resistance to correction. *Nature Reviews Psychology*, 1(1), 13–29. https://doi.org/10.1038/s44159-021-00006-y

Ecker, U. K. H., Roozenbeek, J., Linden, S. Van Der, Tay, L. Q., Cook, J., & Oreskes, N. (2024). Misinformation remains a threat to democracy. *Nature*, 630, 29–32. https://doi.org/10.1038/d41586-024-01587-3

Edwards, S. L., Zarandi, A., Cosimini, M., Chan, T. M., Abudukebier, M., & Stiver, M. L. (2024). Analog serious games for medical education: A scoping review. *Academic Medicine*, 100(3), 375–387. https://doi.org/10.1097/ACM.0000000000005911

Enock, F. E., Bright, J., Stevens, F., Johansson, P., & Margetts, H. Z. (2024). *How do people protect themselves against online misinformation?* The Alan Turing Institute. Retrieved from https://www.turing.ac.uk/news/publications/how-do-people-protect-themselves-against-online-misinformation

Erisen, E., Yildirim, F., Duran, E., Şar, B., & Kalkan, I. (2024). Exploring the effectiveness of virtual reality in combating misinformation on climate change. *Political Psychology*. https://doi.org/10.1111/pops.13057

Erizanu. (2024). Romania's 'rigged' election shows Europe the dangers of Russian disinformation. *The Guardian*. Retrieved from https://www.theguardian.com/commentisfree/2024/dec/11/romania-presidential-election-russia-disinformation-europe

Erlich, A., & Garner, C. (2023). Is pro-Kremlin disinformation effective? Evidence from Ukraine. *International Journal of Press/Politics*, 28(1), 5–28. https://doi.org/10.1177/19401612211045221

European Commission. (2018). *Eurobarometer: Fake news and disinformation online*. European Commission. Retrieved from https://europa.eu/eurobarometer/surveys/detail/2183

Evans, J. S. B. T. (2003). In two minds: Dual-process accounts of reasoning. *Trends in Cognitive Sciences*, 7(10), 454–459. https://doi.org/10.1016/j.tics.2003.08.012

Facebook and Instagram to offer subscription for no ads in Europe. (2024). *Meta Platforms Inc*. Retrieved from https://about.fb.com/news/2024/11/facebook-and-instagram-to-offer-subscription-for-no-ads-in-europe/

Fazio, L. K. (2020). Pausing to consider why a headline is true or false can help reduce the sharing of false news. *Harvard Kennedy School Misinformation Review*. https://doi.org/10.37016/mr-2020-009

Fernando, M., & Arachchilage, N. A. G. (2019). Why Johnny can't rely on anti-phishing educational interventions to protect himself against contemporary phishing attacks? *Australasian Conference on Information Systems*. Retrieved from https://www.cbsnews.com/news/the-phishing-email-that-hacked-the-account-of-john-podesta

Fisher, L., & Andringa, P. (2024). Tories accelerate social media spending in final days of campaign. *Financial Times*. Retrieved from https://www.ft.com/content/645c9f3f-0f3e-4cf0-b230-aca01baef675

Fisher, M., Cox, J. W., & Hermann, P. (2016). Pizzagate: From rumor, to hashtag, to gunfire in D.C. *The Washington Post*. Retrieved from https://www.washingtonpost.com/local/pizzagate-from-rumor-to-hashtag-to-gunfire-in-dc/2016/12/06/4c7def50-bbd4-11e6-94ac-3d324840106c_story.html

Fjællingsdal, K. S., & Klöckner, C. A. (2020). Green Across the Board: Board games as tools for dialogue and simplified environmental communication. *Simulation & Gaming*, 51(5), 632–652. https://doi.org/10.1177/1046878120925133

Flanagan, M., Howe, D. C., & Nissenbaum, H. (2005). Values at play: Design tradeoffs in socially-oriented game design. *CHI 2005*, 751–760. https://doi.org/10.1145/1054972.1055076

Fletcher, R. (2024). Which social networks did political parties use most in 2024? *UK Election Analysis*. Retrieved from https://www.electionanalysis.uk/uk-election-analysis-2024/section-6-the-digital-campaign/which-social-networks-did-political-parties-use-most-in-2024/

Frey, S., Rashid, A., Anthonysamy, P., Pinto-Albuquerque, M., & Naqvi, S. A. (2019). The good, the bad and the ugly: A Study of security decisions in a cyber-physical systems game. *IEEE Transactions on Software Engineering*, 45(5), 521–536. https://doi.org/10.1109/TSE.2017.2782813

Frome, J. (2007). Eight ways videogames generate emotion. *Situated Play, Proceedings of DiGRA 2007 Conference*, 831–835. https://doi.org/10.26503/dl.v2007i1.325

Fuller, I., Edmondson, S., France, D., Higgitt, D., & Ratinen, I. (2006). International perspectives on the effectiveness of geography fieldwork for learning. *Journal of Geography in Higher Education*, 30(1), 89–101. https://doi.org/10.1080/03098260500499667

Fullerton, T. (2004). *Game design workshop: A playcentric approach to creating innovative games* (4th ed.). A K Peters/CRC Press.

Funke, D., & Flamini, D. (2018). A guide to anti-misinformation actions around the world. *The Poynter Institute*. Retrieved from https://www.poynter.org/ifcn/anti-misinformation-actions/

Gavett, B. E., Zhao, R., John, S. E., Bussell, C. A., Roberts, J. R., & Yue, C. (2017). Phishing suspiciousness in older and younger adults: The role of executive functioning. *PLoS ONE*, 12(2), e0171620. https://doi.org/10.1371/journal.pone.0171620

Gee, J. P. (2003). What video games have to teach us about learning and literacy. *Computers in Entertainment*, 1(1), 20–20. https://doi.org/10.1145/950566.950595

Ghadge, A., Weiß, M., Caldwell, N. D., & Wilding, R. (2020). Managing cyber risk in supply chains: A review and research agenda. *Supply Chain Management*, 25(2), 223–240. https://doi.org/10.1108/SCM-10

Gielow Jacobs, L. (2022). Freedom of speech and regulation of fake news. *American Journal of Comparative Law*, 70(Supplement_1), i278–i311. https://doi.org/10.1093/ajcl/avac010

Girard, C., Ecalle, J., & Magnan, A. (2013). Serious games as new educational tools: How effective are they? A meta-analysis of recent studies. *Journal of Computer Assisted Learning*, 29(3), 207–219. https://doi.org/10.1111/j.1365-2729.2012.00489.x

Glickman, H. (2005). The Nigerian "419" advance fee scams: Prank or peril? *Canadian Journal of African Studies / Revue Canadienne Des Études Africaines*, 39(3), 460–489. https://doi.org/10.1080/00083968.2005.10751326

Gondree, M., & Peterson, Z. N. J. (2013). Valuing security by getting [d0x3d!]: Experiences with a network security board game. *6th Workshop on Cyber Security Experimentation and Test, CSET 2013*. Retrieved from https://www.usenix.org/conference/cset13/workshop-program/presentation/gondree

Grace, L., & Hone, B. (2019). Factitious: Large scale computer game to fight fake news and improve news literacy. *Conference on Human Factors in Computing Systems - Proceedings*, 1–8. https://doi.org/10.1145/3290607.3299046

Grace, L., & Liang, S. (2023). Examining misinformation and disinformation games through inoculation theory and transportation theory. *Proceedings of the Annual Hawaii International Conference on System Sciences*, 4691–4700. Retrieved from https://aisel.aisnet.org/hicss-56/ks/game/4/

Grace, L., Dunleavy, V. O., Ahn, R., & Mayo, D. (2024). Designing game based microgames as intervention for health misinformation. *Proceedings of the Annual Hawaii International Conference on System Sciences*, 5513–5521. Retrieved from https://aisel.aisnet.org/hicss-57/ks/security/3/

Graffer, I., Line, M. B., & Bernsmed, K. (2015). *Play2Prepare: A board game supporting it security preparedness exercises for industrial control organizations*. 2015(317631). Retrieved from https://api.semanticscholar.org/CorpusID:55543977

Graham, M. E., Skov, B., Gilson, Z., Heise, C., Fallow, K. M., Mah, E. Y., & Lindsay, D. S. (2023). Mixed news about the bad news game. *Journal of Cognition*, 6(1), 58. https://doi.org/10.5334/joc.324

Graham, T., & Andrejevic, M. (2024). A computational analysis of potential algorithmic bias on platform X during the 2024 US election. Retrieved from https://eprints.qut.edu.au/253211/

Gravel, J., D'Amours-Gravel, M., & Osmanlliu, E. (2023). Learning to fake it: Limited Responses and fabricated references provided by ChatGPT for medical questions. *Mayo Clinic Proceedings: Digital Health*, 1(3), 226–234. https://doi.org/10.1016/j.mcpdig.2023.05.004

Green, M. C., Brock, T. C., & Kaufman, G. F. (2004). Understanding media enjoyment: The role of transportation into narrative worlds. *Communication Theory*, 14(4), 311–327. https://doi.org/10.1111/j.1468-2885.2004.tb00317.x

Green, M., McShane, C. J., & Swinbourne, A. (2022). Active versus passive: Evaluating the effectiveness of inoculation techniques in relation to misinformation about climate change. *Australian Journal of Psychology*, 74(1). https://doi.org/10.1080/00049530.2022.2113340

Grieco, E. (2020). Americans' main sources for political news vary by party and age. *Pew Research Center*. Retrieved from https://www.pewresearch.org/short-reads/2020/04/01/americans-main-sources-for-political-news-vary-by-party-and-age/

Guess, A. M., Lerner, M., Lyons, B., Montgomery, J. M., Nyhan, B., Reifler, J., & Sircar, N. (2020a). A digital media literacy intervention increases discernment between mainstream and false news in the United States and India. *Proceedings of the*

National Academy of Sciences, 117(27), 15536–15545. https://doi.org/10.1073/pnas.1920498117

Guess, A. M., Nyhan, B., & Reifler, J. (2020b). Exposure to untrustworthy websites in the 2016 US election. *Nature Human Behaviour*, 4(5), 472–480. https://doi.org/10.1038/s41562-020-0833-x

Harjani, T., Basol, M.-S., Roozenbeek, J., & van der Linden, S. (2023). Gamified inoculation against misinformation in India: A randomized control trial. *Journal of Trial and Error*, 3(1), 14–56. https://doi.org/10.36850/e12

Harris, E., Rathje, S., Robertson, C. E., & Van Bavel, J. J. (2023). The SPIR framework of social media and polarization: Exploring the role of selection, platform design, incentives, and real-world context. *International Journal of Communication*, 17, 5316–5335. Retrieved from https://ijoc.org/index.php/ijoc/article/view/19014

Hart, S., Margheri, A., Paci, F., & Sassone, V. (2020). Riskio: A serious game for cyber security awareness and education. *Computers and Security*, 95. https://doi.org/10.1016/j.cose.2020.101827

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112. https://doi.org/10.3102/003465430298487

Heerwegh, D. (2009). Mode differences between face-to-face and web surveys: An experimental investigation of data quality and social desirability effects. *International Journal of Public Opinion Research*, 21(1), 111–121. https://doi.org/10.1093/ijpor/edn054

Henderson, K., Napan, K., & Monteiro, S. (2004). Encouraging reflective learning: An online challenge. *Proceedings of the 21st ASCILITE Conference*, 357–364. Retrieved from http://www.ascilite.org.au/conferences/perth04/procs/henderson.html

Henderson, N., & Pallett, H. (2024). *Inoculation theory as a design approach to gamified misinformation interventions*. OSF. https://doi.org/10.31234/osf.io/mr92k

Henderson, N., Buckley, O., & Pallett, H. (2023). Investigating longitudinal effects of physical inoculation interventions against disinformation. In C. Stephanidis, M. Antona, S. Ntoa, & G. Salvendy (Eds.), *HCI international 2023 posters* (pp. 39–46). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-36001-5_5

Henderson, N., Pallett, H., Van Der Linden, S., Montanarini, J., & Buckley, O. (2024). The disPHISHinformation game: Creating a serious game to fight phishing using blended design approaches. *Human Factors in Cybersecurity*, 127, 146–156. https://doi.org/10.54941/ahfe1004774

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

Hill, R. W., Belanich, J., Lane, H. C., Core, M., Dixon, M., Forbell, E., Kim, J., & Hart, J. (2006). Pedagogically structured game-based training: Development of the elect bilat simulation. *Proceedings of the 25th Army Science Conference*. Retrieved from https://apps.dtic.mil/sti/citations/ADA461575

Hochstim, J. R. (1967). A critical comparison of three strategies of collecting data from households. *Journal of the American Statistical Association*, 62(319), 976–989. https://doi.org/10.1080/01621459.1967.10500909

Holliman, R. (2017). Supporting excellence in engaged research. *Journal of Science Communication*, 16(5), 1–10. https://doi.org/10.22323/2.16050304

Hopkins, K. L., Lepage, C., Cook, W., Thomson, A., Abeyesekera, S., Knobler, S., Boehman, N., Thompson, B., Waiswa, P., Ssanyu, J. N., Kabwijamu, L., Wamalwa, B., Aura, C., Rukundo, J. C., & Cook, J. (2023). Co-designing a mobile-based game to improve misinformation resistance and vaccine knowledge in Uganda, Kenya, and Rwanda. *Journal of Health Communication*, 28(sup2), 49–60. https://doi.org/10.1080/10810730.2023.2231377

Hutto, C., & Gilbert, E. (2014). VADER: A parsimonious rule-based model for sentiment analysis of social media text. *Proceedings of the International AAAI Conference on Web and Social Media*, 8(1), 216–225. https://doi.org/10.1609/icwsm.v8i1.14550

International Press Institute. (2020). Hungary: Press freedom threatened as Orbán handed new powers. *International Press Institute*. Retrieved from https://ipi.media/hungary-press-freedom-threatened-as-orban-handed-new-powers/

Ivanov, B. (2012). Designing inoculation messages for health communication campaigns. In Hyunyi. Cho (Ed.), *Health communication message design: Theory and practice* (pp. 73–93). SAGE Publications.

Ivanov, B. (2017). Inoculation theory applied in health and risk messaging. *Oxford Research Encyclopedia of Communication*. https://doi.org/10.1093/acrefore/9780190228613.013.254

Ivanov, B., Miller, C. H., Compton, J., Averbeck, J. M., Harrison, K. J., Sims, J. D., Parker, K. A., & Parker, J. L. (2012). Effects of postinoculation talk on resistance to influence. *Journal of Communication*, 62(4), 701–718. https://doi.org/10.1111/j.1460-2466.2012.01658.x

Ivanov, B., Parker, K. A., & Dillingham, L. L. (2013). Measuring counterargument: A review and critique of the most popular techniques. *The International Journal of Interdisciplinary Studies in Communication*, 7(3), 59–74. https://doi.org/10.18848/2324-7320/cgp/v07i03/53578

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

Ivanov, B., Rains, S. A., Dillingham, L. L., Parker, K. A., Geegan, S. A., & Barbati, J. L. (2022). The role of threat and counterarguing in therapeutic inoculation. *Southern Communication Journal*, 87(1), 15–27. https://doi.org/10.1080/1041794X.2021.1983012

Ivanov, B., Rains, S. A., Geegan, S. A., Vos, S. C., Haarstad, N. D., & Parker, K. A. (2016). Beyond simple inoculation: Examining the persuasive value of inoculation for audiences with initially neutral or opposing attitudes. *Western Journal of Communication*, 81(1), 105–126. https://doi.org/10.1080/10570314.2016.1224917

Jarvis, S., & De Freitas, S. (2009). Evaluation of an immersive learning programme to support triage training. *2009 Conference in Games and Virtual Worlds for Serious Applications*, 117–122. https://doi.org/10.1109/VS-GAMES.2009.31

Jeon, Y., Kim, B., Xiong, A., Lee, D., & Han, K. (2021). ChamberBreaker: Mitigating the echo chamber effect and supporting information hygiene through a gamified inoculation system. *Proceedings of the ACM on Human-Computer Interaction*, *5*(CSCW2). https://doi.org/10.1145/3479859

Johnson, A., & Madsen, J. K. (2024). Inoculation hesitancy: An exploration of challenges in scaling inoculation theory. *Royal Society Open Science*, 11(6). https://doi.org/10.1098/rsos.231711

Johnson, D. W., & Johnson, F. P. (1991). *Joining together: Group theory and group skills*. Prentice-Hall, Inc.

Johnson, D. W., & Johnson, R. T. (1987). *Learning together and alone: Cooperative, competitive, and individualistic learning*. Prentice Hall.

Johnson, D. W., & Johnson, R. T. (2009). An educational psychology success story: Social interdependence theory and cooperative learning. *Educational Researcher*, 38(5), 365–379. https://doi.org/10.3102/0013189X09339057

Johnson, J. R. (1978). The Authenticity and Validity of Antony's Will. *L'Antiquité Classique*, 47(2), 494–503.

Johnson-Glenberg, M. C., Birchfield, D. A., Tolentino, L., & Koziupa, T. (2014). Collaborative embodied learning in mixed reality motion-capture environments: Two science studies. *Journal of Educational Psychology*, 106(1), 86–104. https://doi.org/10.1037/a0034008

Jolley, D., & Paterson, J. L. (2020). Pylons ablaze: Examining the role of 5G COVID-19 conspiracy beliefs and support for violence. *British Journal of Social Psychology*, 59(3), 628–640. https://doi.org/10.1111/bjso.12394

Kahan, D. M., Peters, E., Dawson, E. C., & Slovic, P. (2017). Motivated numeracy and enlightened self-government. *Behavioural Public Policy*, 1(1), 54–86. https://doi.org/10.1017/bpp.2016.2

Kalyuga, S. (2007). Expertise Reversal effect and its implications for learner-tailored instruction. *Educational Psychology Review*, 19, 509–539. https://doi.org/10.1007/s10648-007-9054-3

Kalyuga, S., Ayres, P., Chandler, P., & Sweller, J. (2003). The expertise reversal effect. *Educational Psychologist*, 38(1), 23–31. https://doi.org/10.4018/978-1-60566-048-6.ch003

Karau, S. J., & Williams, K. D. (1993). Social loafing: A meta-analytic review and theoretical integration. *Journal of Personality and Social Psychology*, 65(4), 681–706. https://doi.org/10.1037/0022-3514.65.4.681

Katsaounidou, A., Vrysis, L., Kotsakis, R., Dimoulas, C., & Veglis, A. (2019). MATHE the game: A serious game for education and training in news verification. *Education Sciences*, 9(2). https://doi.org/10.3390/educsci9020155

Kelly Garrett, R., & Weeks, B. E. (2013). The promise and peril of real-time corrections to political misperceptions. *Proceedings of the ACM Conference on Computer Supported Cooperative Work*, CSCW, 1047–1057. https://doi.org/10.1145/2441776.2441895

Keyser, M. W. (2000). Active learning and cooperative learning: Understanding the difference and using both styles effectively. *Research Strategies*, 17(1), 35–44. https://doi.org/10.1016/s0734-3310(00)00022-7

Kiili, K., De Freitas, S., Arnab, S., & Lainema, T. (2012). The design principles for flow experience in educational games. *Procedia Computer Science*, 15, 78–91. https://doi.org/10.1016/j.procs.2012.10.060

Kiili, K., Siuko, J., & Ninaus, M. (2024). Tackling misinformation with games: A systematic literature review. *Interactive Learning Environments*, 32(10), 1–16. https://doi.org/10.1080/10494820.2023.2299999

Kolb, D. A. (1984). *Experiential learning: Experience as the source of learning and development*. Prentice-Hall.

Kolb, D. A., & Fry, R. (1975). Toward an applied theory of experiential learning. In L. Cooper. *Theories of group process* (pp. 33–57). John Wiley & Sons.

Kozyreva, A., Lewandowsky, S., & Hertwig, R. (2020). Citizens versus the internet: Confronting digital challenges with cognitive tools. *Psychological Science in the Public Interest*, 21(3), 103–156. https://doi.org/10.1177/1529100620946707

Kozyreva, A., Lorenz-Spreen, P., Herzog, S. M., Ecker, U. K. H., Lewandowsky, S., Hertwig, R., Ali, A., Bak-Coleman, J., Barzilai, S., Basol, M., Berinsky, A. J., Betsch, C., Cook, J., Fazio, L. K., Geers, M., Guess, A. M., Huang, H., Larreguy, H., Maertens, R., … Wineburg, S. (2024). Toolbox of individual-level interventions against online misinformation. *Nature Human Behaviour*, 8, 1044–1052. https://doi.org/10.1038/s41562-024-01881-0

Kubin, E., & von Sikorski, C. (2021). The role of (social) media in political polarization: A systematic review. *Annals of the International Communication Association*, 45(3), 188–206. https://doi.org/10.1080/23808985.2021.1976070

Kucharski, A. (2016). Post-truth: Study epidemiology of fake news. *Nature*, 540(7634), 525. https://doi.org/10.1038/540525a

Kuklinski, J. H., Quirk, P. J., Jerit, J., Schwieder, D., & Rich, R. F. (2000). Misinformation and the currency of democratic citizenship. *Journal of Politics*, 62(3), 790–816. https://doi.org/10.1111/0022-3816.00033

Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*, 108(3), 480–498. https://doi.org/10.1037/0033-2909.108.3.480

Kyrychenko, Y., Koo, H. J., Maertens, R., Roozenbeek, J., van der Linden, S., & Götz, F. M. (2025). Profiling misinformation susceptibility. *Personality and Individual Differences*, 241. https://doi.org/10.1016/j.paid.2025.113177

Laine, T. H., & Lindberg, R. S. N. (2020). Designing engaging games for education: A systematic literature review on game motivators and design principles. *IEEE Transactions on Learning Technologies*, 13(4), 804–821. https://doi.org/10.1109/TLT.2020.3018503

Lamb, R. L., Annetta, L., Firestone, J., & Etopio, E. (2018). A meta-analysis with examination of moderators of student cognition, affect, and learning outcomes while using serious educational games, serious games, and simulations. *Computers in Human Behavior*, 80, 158–167. https://doi.org/10.1016/j.chb.2017.10.040

Lau, H. M., Smit, J. H., Fleming, T. M., & Riper, H. (2017). Serious games for mental health: Are they accessible, feasible, and effective? A systematic review and meta-analysis. *Frontiers in Psychiatry*, 7(209). https://doi.org/10.3389/fpsyt.2016.00209

Leder, J., Schellinger, L. V., Maertens, R., van der Linden, S., Chryst, B., & Roozenbeek, J. (2024). Feedback exercises boost discernment of misinformation for gamified inoculation interventions. *Journal of Experimental Psychology: General*, 153(8), 2068–2087. https://doi.org/10.1037/xge0001603

Lee, S. K., Sun, J., Jang, S., & Connelly, S. (2022). Misinformation of COVID-19 vaccines and vaccine hesitancy. *Scientific Reports*, 12(1), 13681. https://doi.org/10.1038/s41598-022-17430-6

Lees, J., Banas, J. A., Linvill, D., Meirick, P. C., & Warren, P. (2023). The Spot the Troll Quiz game increases accuracy in discerning between real and inauthentic social media accounts. *PNAS Nexus*, 2(4), 1–11. https://doi.org/10.1093/pnasnexus/pgad094

Leggett, T. (2025). Some JLR suppliers "face bankruptcy" due to hack crisis. *BBC News*. Retrieved from https://www.bbc.co.uk/news/articles/czdjn0lv64ro

Lessig, L. (1999). *Code and other laws of cyberspace*. Basic Books.

Levy, J. (2025). The cyber-driven domino effect: How financial and security crises bankrupt businesses. *MSSP Alert*. Retrieved from https://www.msspalert.com/perspective/the-cyber-driven-domino-effect-how-financial-and-security-crises-bankrupt-businesses

Lewandowsky, S. (2025). Climate change disinformation and how to combat it. *Annu. Rev. Public Health*, 42, 1–21. https://doi.org/annurev-publhealth-090419-102409

Lewandowsky, S., & van der Linden, S. (2021). Countering misinformation and fake news through inoculation and prebunking. *European Review of Social Psychology*, 32(2), 348–384. https://doi.org/10.1080/10463283.2021.1876983

Lewandowsky, S., Ecker, U. K. H., & Cook, J. (2017). Beyond misinformation: Understanding and coping with the "post-truth" era. *Journal of Applied Research in Memory and Cognition*, 6(4), 353–369. https://doi.org/10.1016/j.jarmac.2017.07.008

Lewandowsky, S., Ecker, U. K. H., Seifert, C. M., Schwarz, N., & Cook, J. (2012). Misinformation and its correction: Continued influence and successful debiasing. *Psychological Science in the Public Interest*, 13(3), 106–131. https://doi.org/10.1177/1529100612451018

Lewandowsky, S., Ecker, U. K. H., Cook, J., van der Linden, S., Roozenbeek, J., & Oreskes, N. (2023). Misinformation and the epistemic integrity of democracy. *Current Opinion in Psychology*, *54*. https://doi.org/10.1016/j.copsyc.2023.101711

Lewandowsky, S., Kempe, V., Armaos, K., Hahn, U., Abels, C. M., Wibisono, S., Louis, W., Sah, S., Pagel, C., Jankowicz, N., DiResta, R., Markolin, P., Schoenemann, H., Hertwig, R., Crull, H., Mauer, B., Holford, D., Lopez-Lopez, E., & Cook, J. (2025). *The Anti-Autocracy Handbook A Scholars' Guide to Navigating Democratic Backsliding*. Zenodo. https://doi.org/10.5281/zenodo.15510834

Lewsey, F. (2020). Cambridge game 'pre-bunks' coronavirus conspiracies. *University of Cambridge Stories*. Retrieved from https://www.cam.ac.uk/stories/goviral

Lim, J. S., & Ki, E. J. (2007). Resistance to ethically suspicious parody video on YouTube: A test of inoculation theory. *Journalism and Mass Communication Quarterly*, 84(4), 713–728. https://doi.org/10.1177/107769900708400404

Lofland, John., & Lofland, L. H. (1995). *Analyzing Social Settings: A guide to qualitative observation and analysis* (3rd ed.). Cengage Learning.

Logically. (2024). *Social media, misinformation and harmful algorithms inquiry: written evidence submitted by Logically (SMH0049)*. Retrieved from https://committees.parliament.uk/writtenevidence/132990/html/

Lorant-Royer, S., Munch, C., Mesclé, H., & Lieury, A. (2010). Kawashima vs "Super Mario"! Should a game be serious in order to stimulate cognitive aptitudes? *European Review of Applied Psychology*, 60(4), 221–232. https://doi.org/10.1016/j.erap.2010.06.002

Lorant-Royer, S., Spiess, V., Goncalves, J., & Lieury, A. (2008). Programmes d'entraînement cérébral et performances cognitives: Efficacité, motivation... ou « marketing » ? De la gym-cerveau au programme du dr Kawashima... *Bulletin de Psychologie*, 498(6), 531–549. https://doi.org/10.3917/bupsy.498.0531

Lu, C., Hu, B., Li, Q., Bi, C., & Ju, X. Da. (2023). Psychological inoculation for credibility assessment, sharing intention, and discernment of misinformation: Systematic review and meta-analysis. *Journal of Medical Internet Research*, *25*(e49255). https://doi.org/10.2196/49255

Lucía, A., Zollo, F., Scala, A., Betsch, C., & Quattrociocchi, W. (2018). Polarization of the vaccination debate on Facebook. *Vaccine*, 36(25), 3606–3612. https://doi.org/10.1016/j.vaccine.2018.05.040

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

Ma, J., Chen, Y., Zhu, H., & Gan, Y. (2023). Fighting COVID-19 Misinformation through an online game based on the inoculation theory: Analyzing the mediating effects of perceived threat and persuasion knowledge. *International Journal of Environmental Research and Public Health*, 20(2). https://doi.org/10.3390/ijerph20020980

Maclellan, E. (1999). Reflective commentaries: What do they say about learning? *Educational Action Research*, 7(3), 433–449. https://doi.org/10.1080/09650799900200098

Maertens, R., Götz, F. M., Golino, H. F., Roozenbeek, J., Schneider, C. R., Kyrychenko, Y., Kerr, J. R., Stieger, S., McClanahan, W. P., Drabot, K., He, J., & van der Linden, S. (2024). The Misinformation Susceptibility Test (MIST): A psychometrically validated measure of news veracity discernment. *Behavior Research Methods*, 56(3), 1863–1899. https://doi.org/10.3758/s13428-023-02124-2

Maertens, R., Roozenbeek, J., Basol, M., & van der Linden, S. (2021). Long-term effectiveness of inoculation against misinformation: Three longitudinal experiments. *Journal of Experimental Psychology*: Applied, 27(1), 1–16. https://doi.org/10.1037/xap0000315

Maertens, R., Roozenbeek, J., Simons, J. S., Lewandowsky, S., Maturo, V., Goldberg, B., Xu, R., & van der Linden, S. (2025). Psychological booster shots targeting memory increase long-term resistance against misinformation. *Nature Communications*, 16(1), 2062. https://doi.org/10.1038/s41467-025-57205-x

Maqsood, S., & Chiasson, S. (2021). Design, development, and evaluation of a cybersecurity, privacy, and digital literacy game for tweens. *ACM Transactions on Privacy and Security*, 24(4). https://doi.org/10.1145/3469821

Marsh, T. (2016). Slow serious games, interactions and play: Designing for positive and serious experience and reflection. *Entertainment Computing*, 14, 45–53. https://doi.org/10.1016/j.entcom.2015.10.001

Martel, C., & Rand, D. G. (2024). Fact-checker warning labels are effective even for those who distrust fact-checkers. *Nature Human Behaviour*, 8(10), 1957–1967. https://doi.org/10.1038/s41562-024-01973-x

Mayer, R. E. (2002). Multimedia learning. *Psychology of Learning and Motivation*, 41, 85–139. https://doi.org/10.1016/S0079-7421(02)80005-6

McCosker, A. (2013). Trolling as provocation: YouTube's agonistic publics. *Convergence: The International Journal of Research into New Media Technologies*, 20(2), 201–217. https://doi.org/10.1177/1354856513501413

McGeehan, T. P. (2018). Countering Russian disinformation. *The US Army War College Quarterly: Parameters*, 48(1). https://doi.org/10.55540/0031-1723.2850

McGuire, W. J. (1961). Resistance to persuasion conferred by active and passive prior refutation of the same and alternative counterarguments. *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. *Advances in Experimental Social Psychology*, 1, 191–229. https://doi.org/10.1016/S0065-2601(08)60052-0

McGuire, W. J., & Papageorgis, D. (1961). The relative efficacy of various types of prior belief-defense in producing immunity against persuasion. *Journal of Abnormal and Social Psychology*, 62(2), 327–337. https://doi.org/10.1037/h0042026

McGuire, W. J., & Papageorgis, D. (1962). Effectiveness of forewarning in developing resistance to persuasion. *The Public Opinion Quarterly*, 26(1), 24–34. Retrieved from https://www.jstor.org/stable/2747076

McLoughlin, K. L., Brady, W. J., Goolsbee, A., Kaiser, B., Klonick, K., & Crockett, M. J. (2024). Misinformation exploits outrage to spread online. *Science*, 386(6725), 991–996. https://doi.org/10.1126/science.adl2829

McMahon, L., Kleinman, Z., & Subramanian, C. (2025). Facebook and Instagram get rid of fact checkers. *BBC News*. Retrieved from https://www.bbc.co.uk/news/articles/cly74mpy8klo

Mesa-Gresa, P., Gil-Gómez, H., Lozano-Quilis, J. A., & Gil-Gómez, J. A. (2018). Effectiveness of virtual reality for children and adolescents with autism spectrum disorder: An evidence-based systematic review. *Sensors*, 18(8), 2486. https://doi.org/10.3390/s18082486

Meta Platforms Inc. (2024). Form 10-K. Retrieved from https://www.sec.gov/Archives/edgar/data/1326801/000132680125000017/meta-20241231.htm

Metropolitan Police. (2023). *Cyber Protect: how we can help your business*. Retrieved from https://www.met.police.uk/cyberprotect

Mezirow, J. (1991). *Transformative dimensions of adult learning*. Jossey-Bass

Mezirow, Jack. (1987). *Fostering critical reflection in adulthood: a guide to transformative and emancipatory learning*. Jossey-Bass

Micallef, N., Avram, M., Menczer, F., & Patil, S. (2021). Fakey: A game intervention to improve news literacy on social media. *Proceedings of the ACM on Human-Computer Interaction*, *5*(CSCW1), 1–27. https://doi.org/10.1145/3449080

Miller, J. A., Kosa, M., Gandhi, K., Cooper, S., Iacovides, I., Whitby, M., & Mekler, E. (2024). A design framework for reflective play. *Conference on Human Factors in Computing Systems - Proceedings*. https://doi.org/10.1145/3613904.3642455

Modirrousta-Galian, A., & Higham, P. A. (2023). Gamified inoculation interventions do not improve discrimination between true and fake news: Reanalyzing existing research with receiver operating characteristic analysis. *Journal of Experimental Psychology: General*, 152(9), 2411–2437. https://doi.org/10.1037/xge0001395

Moore, M. (2016). Facebook, the Conservatives and the risk to fair and open elections in the UK. *The Political Quarterly*, 87(3), 424–430. https://doi.org/10.1111/1467-923X.12291

Mora, A., Riera, D., Gonzalez, C., & Arnedo-Moreno, J. (2015). A literature review of gamification design frameworks. *VS-Games 2015 - 7th International Conference on Games and Virtual Worlds for Serious Applications*. https://doi.org/10.1109/VS-GAMES.2015.7295760

Morrison, K. A. (2019). *Study abroad pedagogy, dark tourism, and historical reenactment*. Springer International Publishing. https://doi.org/10.1007/978-3-030-23006-7_3

Morley-Davies, J. (2024). How Did Foreign Actors Exploit the Recent Riots in the UK? *RUSI*. Retrieved from https://www.rusi.org/explore-our-research/publications/commentary/how-did-foreign-actors-exploit-recent-riots-uk

Mousavi, M., Davulcu, H., Ahmadi, M., Axelrod, R., Davis, R., & Atran, S. (2022). Effective messaging on social media: What makes online content go viral? *Proceedings of the ACM Web Conference 2022*, 2957–2966. https://doi.org/10.1145/3485447.3512016

Moyer-Gusé, E., Robinson, M. J., & Mcknight, J. (2018). The role of humor in messaging about the mmr vaccine. *Journal of Health Communication*, 23(6), 514–522. https://doi.org/10.1080/10810730.2018.1473533

National Health Service. (n.d.). *Transformative reflection*. Retrieved from https://www.hee.nhs.uk/our-work/enhancing-generalist-skills/enhance-learning-resources/handbook/cross-cutting-themes/transformative-reflection

Neely, S. R., Eldredge, C., Ersing, R., & Remington, C. (2022). Vaccine hesitancy and exposure to misinformation: A survey analysis. *Journal of General Internal Medicine*, 37(1), 179–187. https://doi.org/10.1007/s11606-021-07171-z

Newman, N., Fletcher, R., Robertson, C. T., Arguedas, A. R., & Nielsen, R. K. (2024). *Digital news report 2024*. Reuters Institute. Retrieved from https://reutersinstitute.politics.ox.ac.uk/digital-news-report/2024

Neylan, J., Biddlestone, M., Roozenbeek, J., & van der Linden, S. (2023). How to "inoculate" against multimodal misinformation: A conceptual replication of Roozenbeek and van der Linden (2020). *Scientific Reports*, 13(1), 1–10. https://doi.org/10.1038/s41598-023-43885-2

Norman, D. A. (2024). *Design for a better world: Meaningful, sustainable, humanity centered*. The MIT Press.

Nosek, B. A., Hardwicke, T. E., Moshontz, H., Allard, A., Corker, K. S., Dreber, A., Fidler, F., Hilgard, J., Struhl, M. K., Nuijten, M. B., Rohrer, J. M., Romero, F., Scheel, A. M., Scherer, L. D., Schönbrodt, F. D., & Vazire, S. (2025). Replicability, robustness, and reproducibility in psychological science. *Annual Review of Psychology*, 16, 51. https://doi.org/10.1146/annurev-psych-122216-011845

Nuñez, F. (2020). Disinformation legislation and freedom of expression. *UC Irvine Law Review*, 10(2), 783–803. Retrieved from https://escholarship.org/uc/item/1x1532sm

Nygren, T., & Guath, M. (2019). Swedish teenagers' difficulties and abilities to determine digital news credibility. *Nordicom Review*, 40(1), 23–42. https://doi.org/10.2478/nor-2019-0002

Nyhan, B., & Reifler, J. (2010). When corrections fail: The persistence of political misperceptions. *Political Behavior*, 32(2), 303–330. https://doi.org/10.1007/s11109-010-9112-2

Oeldorf-Hirsch, A., Schmierbach, M., Appelman, A., & Boyle, M. P. (2020). The ineffectiveness of fact-checking labels on news memes and articles. *Mass Communication and Society*, 23(5), 682–704. https://doi.org/10.1080/15205436.2020.1733613

Ofcom. (2024a). *News consumption in the UK: 2024*. Ofcom. Retrieved from https://www.ofcom.org.uk/media-use-and-attitudes/attitudes-to-news/news-consumption/

Ofcom. (2024b). *Understanding misinformation: an exploration of UK adults' behaviour and attitudes*. Ofcom. Retrieved from https://www.ofcom.org.uk/media-use-and-attitudes/attitudes-to-news/understanding-misinformation-an-exploration-of-uk-adults-behaviour-and-attitudes

Office for National Statistics. (2014). *Public opinions and social trends, Great Britain: 24 April to 6 May 2024*. Office for National Statistics (ONS). Retrieved from https://www.ons.gov.uk/peoplepopulationandcommunity/wellbeing/bulletins/publicopinionsandsocialtrendsgreatbritain/24aprilto6may2024

Oman, P., Schweitzer, E. O., & Roberts, J. (2002). Safeguarding IEDs, substations, and SCADA systems against electronic intrusions. *InterNational Electrical Testing Association 2002 Technical Conference*. Retrieved from https://selinc.com/api/download/2598/

Osmundsen, M., Bor, A., Vahlstrup, P. B., Bechmann, A., & Petersen, M. B. (2021). Partisan polarization is the primary psychological motivation behind political fake news sharing on twitter. *American Political Science Review*, 115(3), 999–1015. https://doi.org/10.1017/S0003055421000290

Online Safety Act 2023, c. 50. Retrieved from https://www.legislation.gov.uk/ukpga/2023/50/enacted

Palmatier, R. W., Houston, M. B., & Hulland, J. (2018). Review articles: Purpose, process, and structure. *Journal of the Academy of Marketing Science*, 46(1), 1–5. https://doi.org/10.1007/s11747-017-0563-4

Pandeirot, L. B., & Aseng, A. C. (2017). Social loafing and group performance: A literature review. *11th International Scholars Conference*, 5(1), 106. https://doi.org/10.35974/isc.v5i1.1655

Paraschivoiu, I., Buchner, J., Praxmarer, R., & Layer-Wagner, T. (2021). Escape the Fake: Development and evaluation of an augmented reality escape room game for fighting fake news. *CHI PLAY 2021 - Extended Abstracts of the 2021 Annual Symposium on Computer-Human Interaction in Play*, 320–325. https://doi.org/10.1145/3450337.3483454

Parker, K. A., Hester, E. B., Ivanov, B., Wombacher, K., & Watterson, T. (2022). Experiencing inoculation: A qualitative analysis of an interrupted message

design. *Western Journal of Communication*, 87(1), 65–85. https://doi.org/10.1080/10570314.2022.2109720

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

Parker, K. A., Ivanov, B., Geegan, S. A., Banas, J. A., & Rains, S. A. (2025). Protective message strategies against binge drinking among college students: Narrative and inoculation messages comparison. *Southern Communication Journal*, 90(2), 101–114. https://doi.org/10.1080/1041794X.2025.2453470

Parker, K. A., Rains, S. A., & Ivanov, B. (2016). Examining the "Blanket of Protection" Conferred by Inoculation: The Effects of inoculation messages on the cross-protection of related attitudes. *Communication Monographs*, 83(1), 49–68. https://doi.org/10.1080/03637751.2015.1030681

Parsons, C., & Drünkler, A. (2019). *Democracy and digital disinformation - How Europe can protect its people without endangering free speech*. Internet Economy Foundation. Retrieved from https://www.ie.foundation/en/blog/ief-demokratie-und-digitale-desinformation

Pearson, M., & Smith, D. (1985). Debriefing in Experience-based Learning. In D. Boud, R. Keogh, D. Walker (Eds.). *Reflection: Turning experience into learning* (pp. 69–84). Routledge.

Pennycook, G., & Rand, D. G. (2019). Fighting misinformation on social media using crowdsourced judgments of news source quality. *Proceedings of the National Academy of Sciences of the United States of America*, 116(7), 2521–2526. https://doi.org/10.1073/pnas.1806781116

Pennycook, G., & Rand, D. G. (2019). Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning. *Cognition*, 188, 39–50. https://doi.org/10.1016/j.cognition.2018.06.011

Pennycook, G., McPhetres, J., Zhang, Y., Lu, J. G., & Rand, D. G. (2020). Fighting COVID-19 misinformation on social media: Experimental evidence for a scalable accuracy-nudge intervention. *Psychological Science*, 31(7), 770–780. https://doi.org/10.1177/0956797620939054

Peppler, K., Danish, J. A., & Phelps, D. (2013). Collaborative gaming: Teaching children about complex systems and collective behavior. *Simulation and Gaming*, 44(5), 683–705. https://doi.org/10.1177/1046878113501462

Perry, C., Slinkard, L. A., Killen, J., & McAlister, A. (1980). Peer teaching and smoking prevention among junior high students. *Adolescence*, 15(58), 277–281. Retrieved from https://www.proquest.com/scholarly-journals/peer-teaching-smoking-prevention-among-junior/docview/1295889702/se-2

Pfau, M., Banas, J., Semmler, S. M., Deatrick, L., Lane, L., Mason, A., Craig, E., Nisbett, G., & Underhill, J. (2010). Role and impact of involvement and enhanced threat in resistance. *Communication Quarterly*, 58(1), 1–18. https://doi.org/10.1080/01463370903520307

Pfau, M., Godbold, L. C., Penaloza, L. J., Yang, V. S., & Hong, Y. (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

Pfau, M., Haigh, M. M., Shannon, T., Tones, T., Mercurio, D., Williams, R., Binstock, B., Diaz, C., Dillard, C., Browne, M., Elder, C., Reed, S., Eggers, A., & Melendez, J. (2008). The influence of television news depictions of the images of war on viewers. *Journal of Broadcasting and Electronic Media*, 52(2), 303–322. https://doi.org/10.1080/08838150801992128

Pfau, M., Ivanov, B., Houston, B., Haigh, M., Sims, J., Gilchrist, E., Russell, J., Wigley, S., Eckstein, J., & Richert, N. (2005). Inoculation and mental processing: The instrumental role of associative networks in the process of resistance to counterattitudinal influence. *Communication Monographs*, 72(4), 414–441. https://doi.org/10.1080/03637750500322578

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, E. A., Anderson, J., Morrill, J., Zubric, J., & 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/10.1093/hcr/27.2.216

Piaget, J. (1972). *Psychology and epistemology: Towards a theory of knowledge*. Penguin Books.

Pomichal, V., & Trnka, A. (2022). Fighting against fake news using the card game "Follow Me." *Proceedings of the European Conference on Games-Based Learning*, 2022-Octob, 455–462. https://doi.org/10.34190/ecgbl.16.1.586

Pomichal, V., & Trnka, A. (2023). With games against fake news – developing critical thinking with the help of the card game Follow Me. *Media Literacy and Academic Research*, 6(1), 55–69. https://doi.org/10.34135/mlar-23-01-04

Quattrociocchi, W., Scala, A., & Sunstein, C. R. (2016). Echo Chambers on Facebook. https://doi.org/10.2139/ssrn.2795110

Ralston, P. A. S., Graham, J. H., & Hieb, J. L. (2007). Cyber security risk assessment for SCADA and DCS networks. *ISA Transactions*, 46(4), 583–594. https://doi.org/10.1016/j.isatra.2007.04.003

Rastenis, J.; Ramanauskaitė, S.; Janulevičius, J.; Čenys, A.; Slotkienė, A.; Pakrijauskas, K. (2020). E-mail-based phishing attack taxonomy. Applied Sciences, 1–15. https://doi.org/10.3390/app10072363

Rathje, S., & van der Linden, S. (2023). Shifting online incentive structures to reduce polarization and the spread of misinformation. In C. R. Sunstein & L. A. Reisch (Eds.), *Research handbook on nudges and society* (pp. 91–108). Edward Elgar Publishing.

Rathje, S., Robertson, C., Brady, W. J., & Van Bavel, J. J. (2023). People think that social media platforms do (but should not) amplify divisive content. *Perspectives on Psychological Science*, 19(5), 781–795. https://doi.org/10.1177/17456916231190392

Raymond S. Nickerson. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, 2(2), 175–220. https://doi.org/10.1037/1089-2680.2.2.175

Rędzio, A. M., Izydorczak, K., Muniak, P., Kulesza, W., & Doliński, D. (2023). Is the COVID-19 bad news game good news? Testing whether creating and disseminating fake news about vaccines in a computer game reduces people's belief in anti-vaccine arguments. *Acta Psychologica*, 236. https://doi.org/10.1016/j.actpsy.2023.103930

Regulation 2022/2065. *Regulation (EU) No 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market for Digital Services and amending Directive 2000/31/EC*. Retrieved from https://eur-lex.europa.eu/eli/reg/2022/2065/oj/eng

Reinheimer, B., Aldag, L., Mayer, P., Mossano, M., Duezguen, R., Lofthouse, B., von Landesberger, T., & Volkamer, M. (2020). An investigation of phishing awareness and education over time: When and how to best remind users. *16th Symposium on Usable Privacy and Security*, 259–284. Retrieved from https://www.usenix.org/conference/soups2020/presentation/reinheimer

Rid, T., & Buchanan, B. (2018). Hacking democracy. *The SAIS Review of International Affairs*, 38(1), 3–16. https://doi.org/10.2307/27001467

Rimé, B., Mesquita, B., Boca, S., & Philippot, P. (1991). Beyond the emotional event: Six studies on the social sharing of emotion. *Cognition and Emotion*, *5*(5–6), 435–465. https://doi.org/10.1080/02699939108411052

Rizvi, S., & Fordham, E. (2025). Cyber security breaches survey 2025.

Roberts, C. (2007). *Mixing modes of data collection in surveys: A methodological review*. National Centre for Research Methods. Retrieved from http://eprints.ncrm.ac.uk/418/1/MethodsReviewPaperNCRM-008.pdf

Robson, K., Plangger, K., Kietzmann, J. H., McCarthy, I., & Pitt, L. (2015). Is it all a game? Understanding the principles of gamification. *Business Horizons*, *5*8(4), 411–420. https://doi.org/10.1016/j.bushor.2015.03.006

Rogerson, M. J., & Gibbs, M. (2018). Finding time for tabletop: Board game play and parenting. *Games and Culture*, 13(3), 280–300. https://doi.org/10.1177/1555412016656324

Rogerson, M. J., Gibbs, M., & Smith, W. (2016). "I love all the bits": The materiality of boardgames. *Conference on Human Factors in Computing Systems - Proceedings*, 3956–3969. https://doi.org/10.1145/2858036.2858433

Roozenbeek, J. (2024). *Propaganda and ideology in the Russian–Ukrainian war*. Cambridge University Press.

Roozenbeek, J., & van der Linden, S. (2018). 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

Roozenbeek, J., & van der Linden, S. (2019). Fake news game confers psychological resistance against online misinformation. *Palgrave Communications*, *5*(1), 65. https://doi.org/10.1057/s41599-019-0279-9

Roozenbeek, J., & van der Linden, S. (2020). Breaking Harmony Square: A game that "inoculates" against political misinformation. *Harvard Kennedy School Misinformation Review*. https://doi.org/10.37016/mr-2020-47

Roozenbeek, J., & van der Linden, S. (2024). *The psychology of misinformation*. Cambridge University Press.

Roozenbeek, J., Culloty, E., & Suiter, J. (2023). Countering misinformation: Evidence, knowledge gaps, and implications of current interventions. *European Psychologist*, 28(3), 189–205. https://doi.org/10.1027/1016-9040/a000492

Roozenbeek, J., Maertens, R., McClanahan, W., & van der Linden, S. (2021). Disentangling item and testing effects in inoculation research on online misinformation: Solomon revisited. *Educational and Psychological Measurement*, 81(2), 340–362. https://doi.org/10.1177/0013164420940378

Roozenbeek, J., Remshard, M., & Kyrychenko, Y. (2024). Beyond the headlines : On the efficacy and effectiveness of misinformation interventions. *Advances.in/Psychology*, 2, e24569. https://doi.org/10.56296/aip00019

Roozenbeek, J., Schneider, C. R., Dryhurst, S., Kerr, J., Freeman, A. L. J., Recchia, G., Van Der Bles, A. M., & Van Der Linden, S. (2020). Susceptibility to misinformation about COVID-19 around the world. *Royal Society Open Science*, 7(10), 201199. https://doi.org/10.1098/rsos.201199

Roozenbeek, J., Traberg, C. S., & van der Linden, S. (2022). Technique-based inoculation against real-world misinformation. *Royal Society Open Science*, 9, 211719. https://doi.org/10.1098/rsos.211719

Roozenbeek, J., van der Linden, S., & Lewandowsky, S. (2022). YouTube: How a team of scientists worked to inoculate a million users against misinformation. *The Conversation*. Retrieved from https://theconversation.com/youtube-how-a-team-of-scientists-worked-to-inoculate-a-million-users-against-misinformation-189007

Roozenbeek, J., van der Linden, S., & Nygren, T. (2020). Prebunking interventions based on "inoculation" theory can reduce susceptibility to misinformation across cultures. *Harvard Kennedy School Misinformation Review*, 1(2). https://doi.org/10.37016//mr-2020-008

Roy, A., G, N., & Mukherjee, S. (2025). *ShieldUp!: Inoculating users against online scams using a game based intervention*. arXiv. Retrieved from http://arxiv.org/abs/2503.12341

Saleh, N. F., Makki, F., van der Linden, S., & Roozenbeek, J. (2023). Inoculating against extremist persuasion techniques – Results from a randomised controlled trial in post-conflict areas in Iraq. *Advances.in/Psychology*, 1, 1–22. https://doi.org/10.56296/aip00005

Saleh, N. F., Roozenbeek, J. O. N., Makki, F. A., Mcclanahan, W. P., & Van Der Linden, S. (2024). Active inoculation boosts attitudinal resistance against extremist persuasion techniques: a novel approach towards the prevention of violent extremism. *Behavioural Public Policy*, 8(3), 548–571. https://doi.org/10.1017/bpp.2020.60

Salen, K., & Zimmerman, E. (2004). *Rules of play: Game design fundamentals*. MIT Press.

Sarno, D. M., & Black, J. (2024). Who gets caught in the web of lies?: Understanding susceptibility to phishing emails, fake news headlines, and scam text messages. *Human Factors*, 66(6), 1742–1753. https://doi.org/10.1177/00187208231173263

Sarno, D. M., Lewis, J. E., Bohil, C. J., & Neider, M. B. (2020). Which phish is on the hook? Phishing vulnerability for older versus younger adults. *Human Factors*, 62(5), 704–717. https://doi.org/10.1177/0018720819855570

Sarno, D. M., University, C., South, C., & Neider, M. B. (2021). So many phish, so little time: Exploring email task factors and phishing susceptibility. *Human factors*, 64(8), 1379–1403. https://doi.org/10.1177/0018720821999174

Sasaki, S., Kurokawa, H., & Ohtake, F. (2021). Effective but fragile? Responses to repeated nudge-based messages for preventing the spread of COVID-19 infection. *The Japanese Economic Review*, 72(3). https://doi.org/10.1007/s42973-021-00076-w

Scanlan, J. M., & Chernomas, W. M. (1997). Developing the reflective teacher. *Journal of Advanced Nursing*, 25(6), 1138–1143. https://doi.org/10.1046/j.1365-2648.1997.19970251138.x

Schell, J. (2019). *The art of game design: A book of lenses* (3rd Edition). A K Peters/CRC Press.

Schmidt, S. (2009). Shall we really do it again? The powerful concept of replication is neglected in the social sciences. *Review of General Psychology*, 13(2), 90–100. https://doi.org/10.1037/a0015108

Schön, D. A. (1984). *The reflective practitioner: How professionals think in action*. Basic Books.

Sethi, R. J. (2017). Crowdsourcing the verification of fake news and alternative facts. *Proceedings of the 28th ACM Conference on Hypertext and Social Media*, 44(8), 315–316. https://doi.org/10.1088/1751-8113/44/8/085201

Shead, S. (2022). Meta says it may shut down Facebook and Instagram in Europe over data-sharing dispute. *CNBC*. Retrieved from https://www.cnbc.com/2022/02/07/meta-threatens-to-shut-down-facebook-and-instagram-in-europe.html

Shepard, R. N. (1967). Recognition memory for words, sentences, and pictures. *Journal of Verbal Learning and Verbal Behavior*, 6(1), 156–163. https://doi.org/10.1016/S0022-5371(67)80067-7

Silverman, C. (2025). As Facebook Abandons Fact-Checking, It's Also Offering Bonuses for Viral Content. *ProPublica*. Retrieved from https://www.propublica.org/article/facebook-meta-abandons-fact-checking-boosts-viral-content

Sirianni, F. A. (1984). Was Antony's will partially forged? *L'Antiquité Classique*, 53, 236–241. https://doi.org/10.3406/antiq.1984.2126

Smith, E., & Golding, L. (2018). Use of board games in higher education literature review. *MSOR Connections*, 16(2), 24–29. https://doi.org/10.21100/msor.v16i2.624

Standing, L. (1973). Learning 10000 pictures. *Quarterly Journal of Experimental Psychology*, 25(2), 207–222. https://doi.org/10.1080/14640747308400340

Stanfield, M., & Connolly, T. (2008). Cognitive and affective effects of learning history by playing a mobile game. 2n*d European Conference on Games-Based Learning (ECGBL 2008)*, 207–212. Retrieved from https://dare.uva.nl/search?identifier=3aaeae73-9af5-44ea-8687-83b9651736fd

State of Agile. (2023). *17th state of agile report*. Digital.AI. Retrieved from https://digital.ai/resource-center/analyst-reports/state-of-agile-report/

Stephan, W. G. (2014). Intergroup anxiety: Theory, research, and practice. *Personality and Social Psychology Review*, 18(3), 239–255. https://doi.org/10.1177/1088868314530518

Stephan, W. G., & Stephan, C. W. (1985). Intergroup anxiety. *Journal of Social Issues*, 41(3), 157–175. https://doi.org/10.1111/j.1540-4560.1985.tb01134.x

Stepney, E. S., & Lally, C. (2024). *Disinformation: sources, spread and impact*. UK Parliament. https://doi.org/10.58248/PN719

Swami, V., Voracek, M., Stieger, S., Tran, U. S., & Furnham, A. (2014). Analytic thinking reduces belief in conspiracy theories. *Cognition*, 133(3), 572–585. https://doi.org/10.1016/j.cognition.2014.08.006

Tandoc, E., & Seet, S. (2023). Winning the game against fake news? Using games to inoculate adolescents and young adults in Singapore against fake news. *Estudios Sobre El Mensaje Periodistico*, 29(4), 771–781. https://doi.org/10.5209/esmp.88599

Taspinar, B., Schmidt, W., & Schuhbauer, H. (2016). Gamification in education: A board game approach to knowledge acquisition. *Procedia Computer Science*, 99, 101–116. https://doi.org/10.1016/j.procs.2016.09.104

Thorne, J., & Vlachos, A. (2018). Automated fact checking: Task formulations, methods and future directions. In E. M. Bender, L. Derczynski, & P. Isabelle (Eds.),

COLING 2018 - 27th International conference on computational linguistics, proceedings (pp. 3346–3359). Association for Computational Linguistics.

Traberg, C. S., & van der Linden, S. (2022). Birds of a feather are persuaded together: Perceived source credibility mediates the effect of political bias on misinformation susceptibility. *Personality and Individual Differences*, 185, 111269. https://doi.org/10.1016/j.paid.2021.111269

Traberg, C. S., Roozenbeek, J., & van der Linden, S. (2022). Psychological inoculation against misinformation: Current evidence and future directions. *Annals of the American Academy of Political and Social Science*, 700(1), 136–151. https://doi.org/10.1177/00027162221087936

Traberg, C. S., Roozenbeek, J., & van der Linden, S. (2024). Gamified inoculation reduces susceptibility to misinformation from political ingroups. *Harvard Kennedy School Misinformation Review*, *5*(2). https://doi.org/10.37016/mr-2020-141

Threat hunting report (2025). *CrowdStrike 2025 threat hunting report*. CrowdStrike. Retrieved from https://go.crowdstrike.com/2025-threat-hunting-report.html

UK Department for Education. (2024). *Government launches curriculum and assessment review*. Department for Education (UK Government). Retrieved from https://www.gov.uk/government/news/government-launches-curriculum-and-assessment-review

UK National Cyber Security Centre. (2017). *Guidance; Small Business Guide: Cyber Security; Step 5 - Avoiding phishing attacks*. Retrieved from https://www.ncsc.gov.uk/collection/small-business-guide/avoiding-phishing-attacks

Ulven, J. B., & Wangen, G. (2021). A systematic review of cybersecurity risks in higher education. *Future Internet*, 13(39), 1–40. https://doi.org/10.3390/fi13020039

Valenzuela, S., Piña, M., & Ramírez, J. (2017). Behavioral effects of framing on social media users: How conflict, economic, human interest, and morality frames drive news sharing. *Journal of Communication*, 67(5), 803–826. https://doi.org/10.1111/jcom.12325

Van Bavel, J. J., Harris, E. A., Pärnamets, P., Rathje, S., Doell, K. C., & Tucker, J. A. (2021). Political Psychology in the Digital (mis)Information age: A model of news belief and sharing. *Social Issues and Policy Review*, 15(1), 84–113. https://doi.org/10.1111/sipr.12077

van der Linden, S., Roozenbeek, J., Maertens, R., Basol, M., Kácha, O., Rathje, S., & Traberg, C. S. (2021). How can psychological science help counter the spread of

fake news? *Spanish Journal of Psychology*, 24(e25). https://doi.org/10.1017/SJP.2021.23

van der Linden, S. (2022). Misinformation: susceptibility, spread, and interventions to immunize the public. *Nature Medicine*, 28(3), 460–467. https://doi.org/10.1038/s41591-022-01713-6

van der Linden, S., & Kyrychenko, Y. (2024). A broader view of misinformation reveals potential for intervention. *Science*, 384(6699), 959. https://doi.org/10.1126/science.adp9117

van der Linden, S., & Roozenbeek, J. (2020). Psychological inoculation against fake news. In R. Greifeneder, M. Jaffe, E. Newman, & N. Schwarz (Eds.), *The psychology of fake news: Accepting, sharing, and correcting misinformation* (pp. 147–169). Routledge. https://doi.org/10.4324/9780429295379-11

van der Linden, S., Roozenbeek, J., Maertens, R., Basol, M., Kácha, O., Rathje, S., & Traberg, C. S. (2021). How can psychological science help counter the spread of fake news? *Spanish Journal of Psychology*, 24(e25). https://doi.org/10.1017/SJP.2021.23

Vicario, M. Del, Bessi, A., Zollo, F., Petroni, F., Scala, A., Caldarelli, G., Stanley, H. E., & Quattrociocchi, W. (2016). The spreading of misinformation online. *Proceedings of the National Academy of Sciences of the United States of America*, 113(3), 554–559. https://doi.org/10.1073/pnas.1517441113

Vraga, E. K., Kim, S. C., & Cook, J. (2019). Testing logic-based and humor-based corrections for science, health, and political misinformation on social media. *Journal of Broadcasting and Electronic Media*, 63(3), 393–414. https://doi.org/10.1080/08838151.2019.1653102

Walker, M., & Matsa, K. E. (2021). *News consumption across social media in 2021: More than half of twitter users get news on the site regularly*. Pew Research Center. Retrieved from http://www.jstor.org/stable/resrep63516

Wallack, L., & Corbett, K. (1987). Alcohol, tobacco and marijuana use among youth: An overview of epidemiological, program and policy trends. *Health Education Quarterly*, 14(2), 223–249. https://doi.org/10.1177/109019818701400208

Walter, N., Cohen, J., Holbert, R. L., & Morag, Y. (2020). Fact-checking: A meta-analysis of what works and for whom. *Political Communication*, 37(3), 350–375. https://doi.org/10.1080/10584609.2019.1668894

Wardle, C., & Derakhshan, H. (2017). *Information disorder: Toward an interdisciplinary framework for research and policymaking*. Council of Europe. Retrieved from

https://edoc.coe.int/en/media/7495-information-disorder-toward-an-interdisciplinary-framework-for-research-and-policy-making.html

Warren, J. (2025). Ransomware attack contributed to patient's death. *BBC News*. Retrieved from https://www.bbc.co.uk/news/articles/cp3ly4v2kp2o

Watson, J., van der Linden, S., Watson, M., & Stillwell, D. (2024). Negative online news articles are shared more to social media. *Scientific Reports*, 14(1), 21592. https://doi.org/10.1038/s41598-024-71263-z

Wendel, V., Gutjahr, M., Göbel, S., & Steinmetz, R. (2013). Designing collaborative multiplayer serious games. *Education and Information Technologies*, 18(2), 287–308. https://doi.org/10.1007/s10639-012-9244-6

Whittemore, R., & Knafl, K. (2005). The integrative review: Updated methodology. *Journal of Advanced Nursing*, *52*(5), *546–553*. https://doi.org/10.1111/j.1365-2648.2005.03621.x

Williams, J. (2018). *Stand out of our light: Freedom and resistance in the attention economy*. Cambridge University Press.

Williams, J. H., Hooker, C., Gilbert, G. L., Hor, S., & Degeling, C. (2023). Disagreement among experts about public health decision making: Is it polarisation and does it matter? *BMJ Global Health*, 8(3). https://doi.org/10.1136/bmjgh-2022-011182

Williams, P., & Sheridan, S. (2010). Conditions for collaborative learning and constructive competition in school. *Educational Research*, *52*(4), 335–350. https://doi.org/10.1080/00131881.2010.524748

Wimmer, G. E., & Poldrack, R. A. (2021). Reward learning and working memory: Effects of massed versus spaced training and post-learning delay period. *Memory & Cognition*, *50*(2), 312–324. https://doi.org/0.3758/s13421-021-01233-7

Wineburg, S., Breakstone, J., Smith, M., McGrew, S., & Ortega, T. (2019). *Civic Online Reasoning: Curriculum Evaluation*. Civic Online Reasoning. Retrieved from https://purl.stanford.edu/xr124mv4805

Wong, C. M. L., & Wu, Y. (2023). Limits to inoculating against the risk of fake news: a replication study in Singapore during COVID-19. *Journal of Risk Research*, 26(10), 1037–1052. https://doi.org/10.1080/13669877.2023.2249909

Wood, M. L. M. (2007). Rethinking the inoculation analogy: Effects on subjects with differing preexisting attitudes. *Human Communication Research*, 33(3), 357–378. https://doi.org/10.1111/j.1468-2958.2007.00303.x

Wood, Z. (2025). 'Hello pervert': The sextortion scam claiming to have videoed you. *The Guardian*. Retrieved from https://www.theguardian.com/money/2025/may/11/hello-pervert-sextortion-scam-video-email-hacked

World Health Organisation. (2020). *Countering misinformation about COVID-19*. World Health Organisation. Retrieved from https://www.who.int/news-room/feature-stories/detail/countering-misinformation-about-covid-19

Wouters, P., & Van Oostendorp, H. (2013). A meta-analytic review of the role of instructional support in game-based learning. *Computers and Education*, 60(1), 412–425. https://doi.org/10.1016/j.compedu.2012.07.018

Wouters, P., van Nimwegen, C., van Oostendorp, H., & van Der Spek, E. D. (2013). A meta-analysis of the cognitive and motivational effects of serious games. *Journal of Educational Psychology*, 105(2), 249–265. https://doi.org/10.1037/a0031311

Wu, D., Zhang, J., Brown, N., Lowry, P. B., & Moody, G. D. (2020). Patching the "human" in information security: Using the inoculation defense to confer resistance against phishing attacks. *Annual Pre-ICIS Workshop on HCI Research in MIS 2020*. Retrieved from https://aisel.aisnet.org/sighci2020

Xie, L., Antle, A. N., & Motamedi, N. (2008). Are tangibles more fun? Comparing children's enjoyment and engagement using physical, graphical and tangible user interfaces. *Proceedings of the 2nd International Conference on Tangible and Embedded Interaction*, 191–198. https://doi.org/10.1145/1347390.1347433

Yaman, M., Nerdel, C., & Bayrhuber, H. (2008). The effects of instructional support and learner interests when learning using computer simulations. *Computers and Education*, 51(4), 1784–1794. https://doi.org/10.1016/j.compedu.2008.05.009

Yaqub, W., Kakhidze, O., Brockman, M. L., Memon, N., & Patil, S. (2020). Effects of Credibility Indicators on Social Media News Sharing Intent. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–14. https://doi.org/10.1145/3313831.3376213

Zagal, J. P., Rick, J., & Hsi, I. (2006). Collaborative games: Lessons learned from board games. *Simulation and Gaming*, 37(1), 24–40. https://doi.org/10.1177/1046878105282279

Zeichner, K. M. (1994). Research on Teacher Thinking and Different Views of Reflective Practice in Teaching and Teacher Education. In I. Carlgren, G. Handal, S. Vaage (Eds.). *Teachers' minds and actions: Research on teachers' thinking and practice* (pp. 9–27). Routledge.

Zerback, T., Töpfl, F., & Knöpfle, M. (2021). The disconcerting potential of online disinformation: Persuasive effects of astroturfing comments and three strategies for inoculation against them. *New Media and Society*, 23(5), 1080–1098. https://doi.org/10.1177/1461444820908530

Zhonggen, Y. (2019). A meta-analysis of use of serious games in education over a decade. *International Journal of Computer Games Technology*, 2019(1), 4797032. https://doi.org/10.1155/2019/4797032

10. Appendices

Appendix 1: Study 1 Variable Correlation Matrix

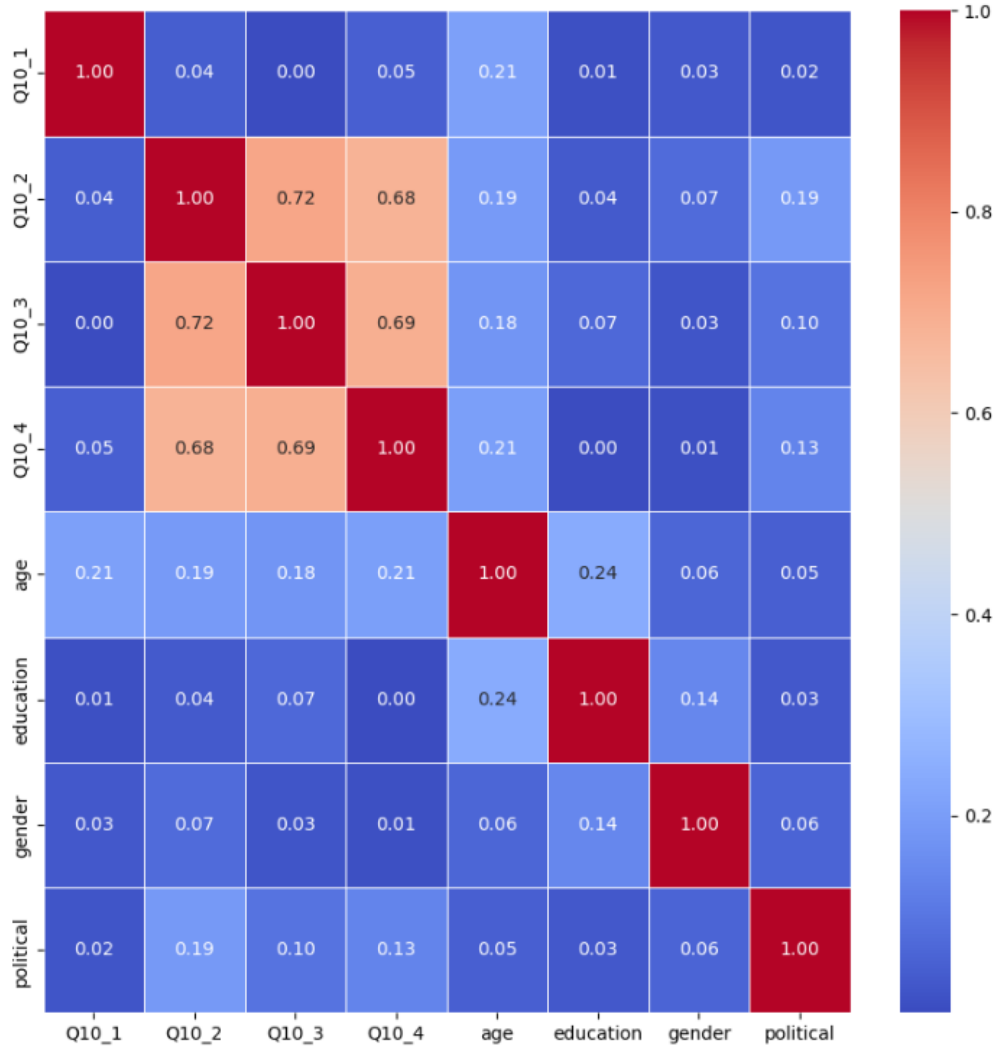


Figure: Bivariate (Pearson) correlation matrix of variables in Chapter 4 (Q10_1=issue familiarity; Q10_2=perceived persuasiveness; Q10_3=agreement; Q10_4=perceived reliability).

Appendix 2: Main Statistical Tests from Study 1

Table: Results from Kruskal-Wallis test with given variables between T1, T2, and T3.

Variable	Df	H statistic	*p*
Persuasiveness	2	1.25	.53
Agreement	2	0.52	.77
Reliability	2	1.09	.58

Table: One-tailed independent T-Test results on given variables between T3 and online control.

Variable	Df	T statistic	*p*
Persuasiveness	138	-0.73	.23
Agreement	138	0.32	.62
Reliability	138	-0.65	.26

Appendix 3: Moderation Analysis of Issue Familiarity in Study 1

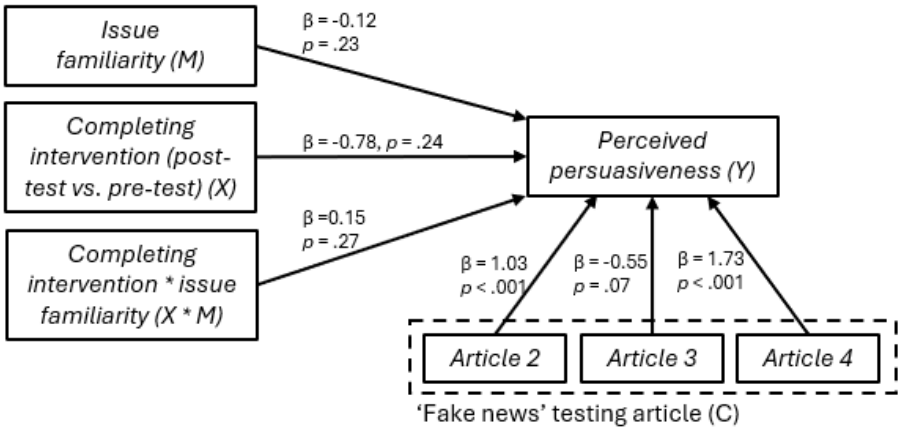


Figure: Moderation analyses of issue familiarity for intervention's effect on perceived persuasiveness of 'fake news' articles (R^2 = .23, F(6, 235) = 11.15, p < .001).

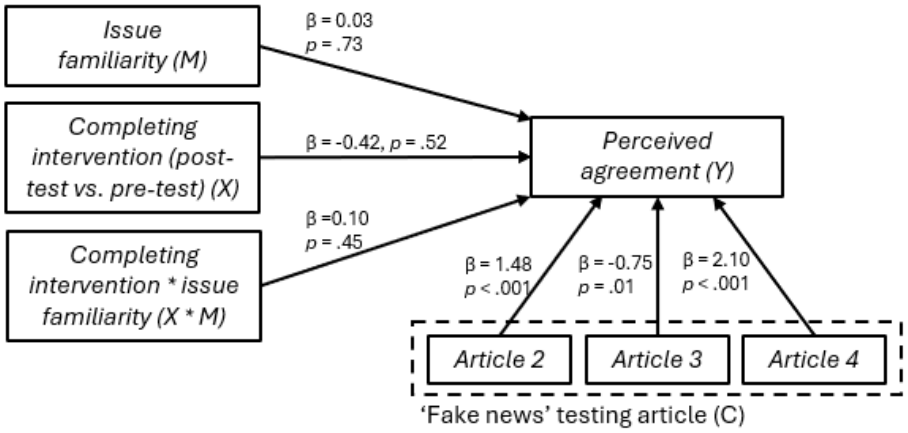


Figure: Moderation analyses of issue familiarity for intervention's effect on agreement with 'fake news' articles (R^2 = .33, F(6, 235) = 18.37, p < .001).

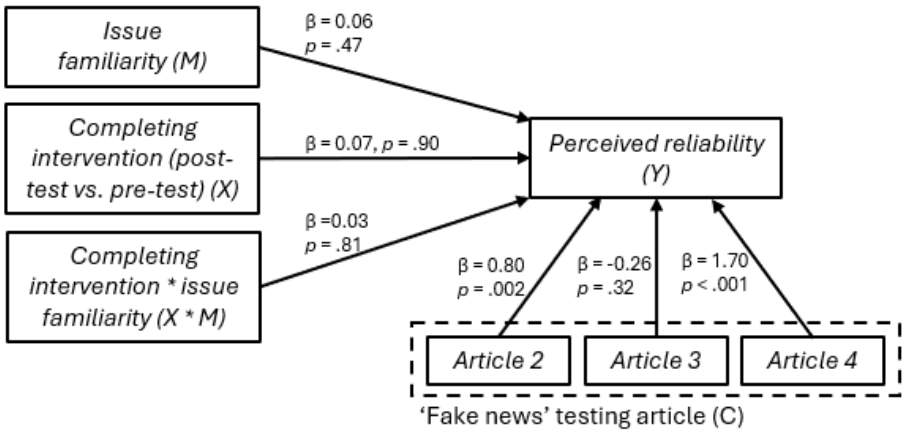


Figure: Moderation analyses of issue familiarity for intervention's effect on perceived reliability of 'fake news' articles (R^2 = .23, F(6, 235) = 11.14, p < .001).

Appendix 4: Statistical Tests with Socio-demographic Variables in Study 1

Table: Pearson Correlation Coefficient between individuals' political leaning and three given variables.

Variable	Df	Correlation statistic	*p*
Persuasiveness	327	0.19	<.001
Agreement	327	0.10	.08
Reliability	327	0.13	.02

Table: Pearson Correlation Coefficient between age and three given variables.

Variable	Df	Correlation statistic	*p*
Persuasiveness	327	-0.19	< .001
Agreement	327	-0.18	.001
Reliability	327	-0.21	< .001

Table: Pearson Correlation Coefficient between education and three given variables.

Variable	Df	Correlation statistic	*p*
Persuasiveness	327	-0.039	.48
Agreement	327	-0.068	.22
Reliability	327	-0.005	.93

Appendix 5: Moderation Analysis of Education in Study 1

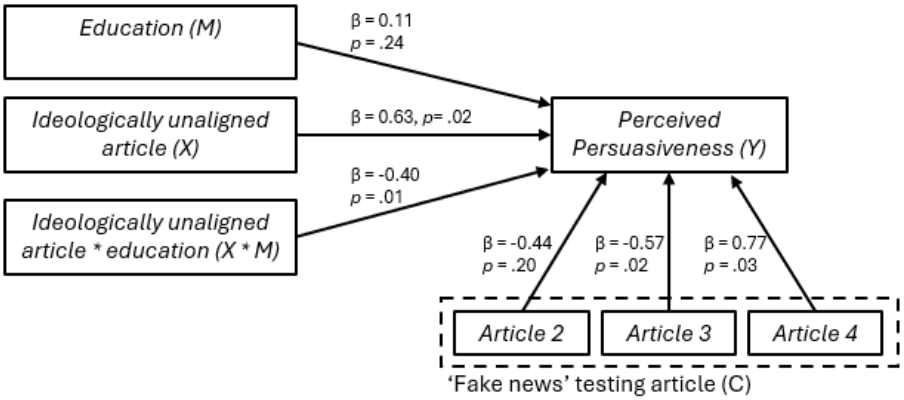


Figure: Moderating effect of education on how the ideological leaning of the 'fake news' article (in relation to the participant) affects its perceived persuasiveness (R^2 = .24, F(6, 273) = 14.02, p < .001).

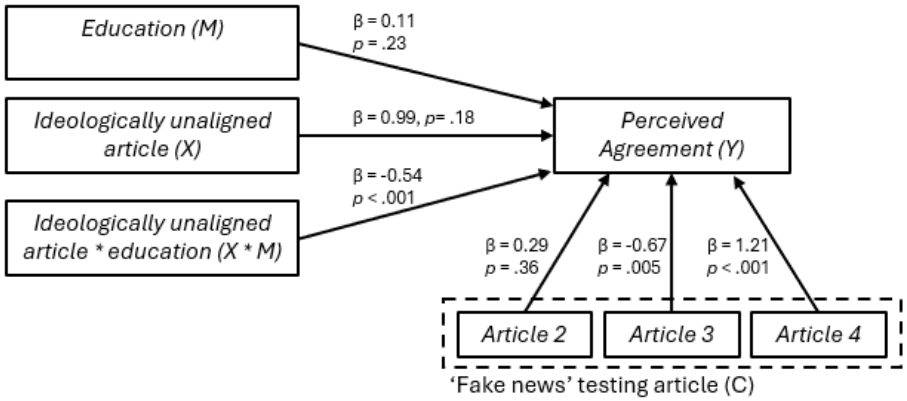


Figure: Moderating effect of education on how the ideological leaning of the 'fake news' article (in relation to the participant) effects how much participants agree with it (R^2 = .42, F(6, 274) = 32.4, p < .001).

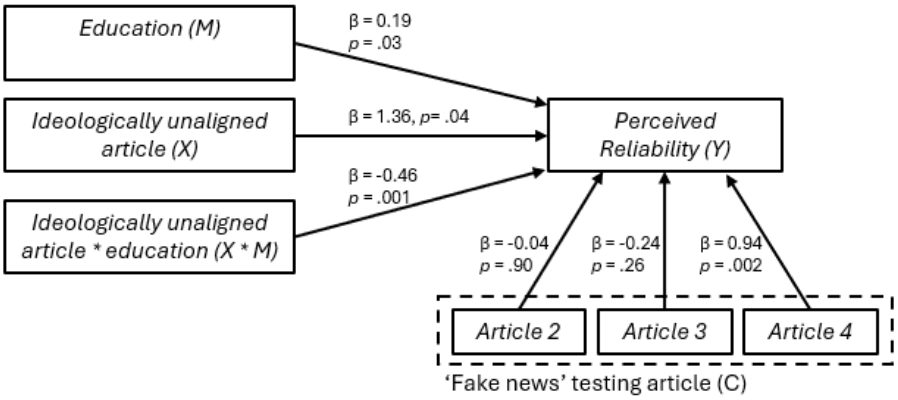


Figure: Moderating effect of education on how the ideological leaning of the 'fake news' article (in relation to the participant) affects its perceived persuasiveness (R^2 = .21, F(6, 273) = 12.14, p < .001).

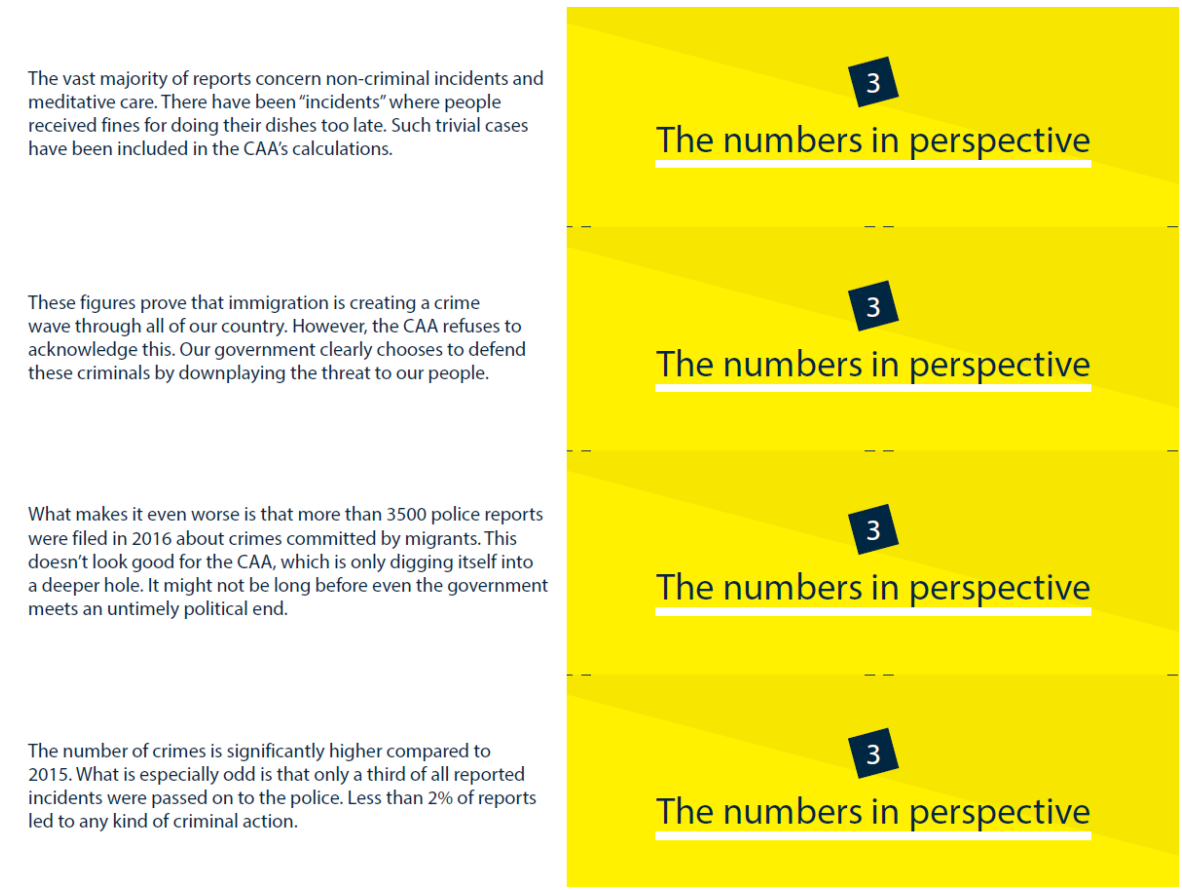


Figure: Example cards for the game-based intervention Fake News. Pictured, 'the numbers in perspective' article section.

Figure: Example 'source' card from game-based intervention Fake News.

Figure: Two 'character' cards from the game-based intervention Fake News.

Appendix 7: Study 1 Reading Task

Table: Four 'fake news' articles used in the thesis' first study (see Chapter 4 and 5).

Article 1, right leaning
Shocking facts: "Up to 100,000" migrants on their way to the United Kingdom. *Hordes of migrants in Libya are once again aching to cross over to Europe. The British government is ignoring the problem and stubbornly maintains that nothing is wrong. The British are worried.* **Huge influx** According to new research by the Centre for Political Analysis, up to 100.000 fortune seekers are anxiously waiting in Libya to make their way to Europe. This is a shocking rise compared to last year, when the influx was already very high. **Europe's "plan"** The European Union has met the new report with rejection. Jean-Claude Juncker, President of the European Commission, called it "total nonsense", even though everybody knows that the number of immigrants at the gates of Europe has risen precipitously. Juncker justifies his stance by saying that all refugees deserve a safe haven. But in the meantime he does nothing for European pensioners: for years, they've felt more and more insecure in their financial situation and are afraid to lose their income, and Europe doesn't do anything about it. There is of course always the possibility that Juncker had had one too many when he was reading the report. **Is something being swept under the rug?** Every British person knows very well that the number of fortune seekers wanting to enter our country through European channels is simply too large to be sustainable. This is also acknowledged by independent experts. "Juncker has a strong interest in sweeping the real immigration numbers under the rug", says Nelly Dowling, a respected immigration expert. "If he'd be telling the truth, all of Europe would call for his resignation. So he pretends like nothing is wrong." And in the meantime, the United Kingdom is losing its faith in Europe more and more.
Article 2, right leaning
Report: refugee camps "just like in World War II" *The situation in East European refugee camps is a total disaster, says a new report. The European Union does nothing to save refugees from a sad fate. British people and Europeans look upon it with sorrow.* **Concentration camps** Researchers at the Centre for Reliable Statistics published a report last week in which the shocking situation in East European refugee camps was exposed. Up to a million people are locked up in small, dirty cages without clean drinking water or enough food. This is especially prevalent in countries like Hungary, Greece, Serbia and Slovakia. **What does the EU do?** The European Union says that the situation "warrants attention" but is "under control". Dmitris Avramopoulos, the EU Commissioner for Migration, Home Affairs and Citizenship, stated that he was not impressed by the new numbers. The images, however, speak for themselves: everyone with working retinas can see that it can't go on like this. Avramopoulos defends himself by pointing out that every European country is responsible for its own refugee policy. But at the same time he also thinks the EU should decide how much parental leave new fathers should be allowed to get. This is, of course, rank hypocrisy.

Avramopoulos (who by the way looks a bit like the evil Emperor Palpatine from Star Wars) happens to be from Greece, which is one of the countries where refugee camps are far below par. The possibility that he is trying to cover for his own government by sweeping the truth under the carpet shouldn't be excluded.

Experts: the situation is hopeless
Every sane British and European will understand that the European Union cannot simply abandon these people. Many serious experts agree. "The situation is hopeless. Under pressure from rightwing populists, Europe is doing nothing. It's almost as if Europe has struck a backroom deal with East European leaders who are against immigration, like Viktor Órban. Of course this is being kept hidden from the general public", says Henk Oostra, a world-famous scientist and statistician. Our human values are at stake.

Article 3, left leaning

Report: British culture being watered down by European immigration
British culture is disappearing in the face of the hordes of immigrants currently invading Britain, says a new report. The feather soft touch of the government has allowed increasingly high levels of immigration, reducing our British identity in place of foreign restaurants, supermarkets, and bars.

We've all seen it: the disappearing pubs
The British pub, an establishment that has quenched thirsts from the medieval era to the industrial revolution, is slowly disappearing in UK towns along with other shops such as butchers, bakers, and fishmongers, according to the new study. The number of pubs that operate in the UK is now 30% less than in 2020. This is a worrying figure, and if you enjoy a drink after a long working Friday, you should be worried.

The immigration problem
A new report by the Citizen's Centre for British Culture and Tourism (CCBCT) has blamed this on "chaos at the borders". The report claims "the unchecked immigration from European countries such as Hungary, Poland, and Romania has watered down British buildings and British values. Selfish immigrants from Hungary and Poland are intentionally spelling the end of what it is to be British".

Charles Barnaby, a leading expert in British culture, has suggested that leaders at the top need to rule with an iron fist: "Those who campaigned to us get out of the European Union understood the dangers of this unchecked immigration. Current leaders are turning a blind eye!" Our current home secretary, Priti Patel, claims to have a tough approach to immigration. Maybe she needs glasses – she may be better than the foreign-loving left, but not tough enough! Everyone knows that what this country needs, is to be by the British, for the British.

CCBCT Report: Next Steps
The CCBCT report has outlines some next steps in maintaining hold of our British culture, but we all know that it starts with leadership change for a political party more prepared to tackle this rising immigration. And until that happens, we all see our British culture slowly slips away.

Article 4, right leaning

Immoral: Relatives seeking asylum in the United Kingdom unable to join settled families after immigration U-turn.
Refugees are being forced to return home after grueling small boat crossings over the English Channel, even if they currently have family in the United Kingdom. The British government's sudden shrewd U-turn has visibly put the lives of many at risk.

We can all see it: government decision cuts deep for many communities

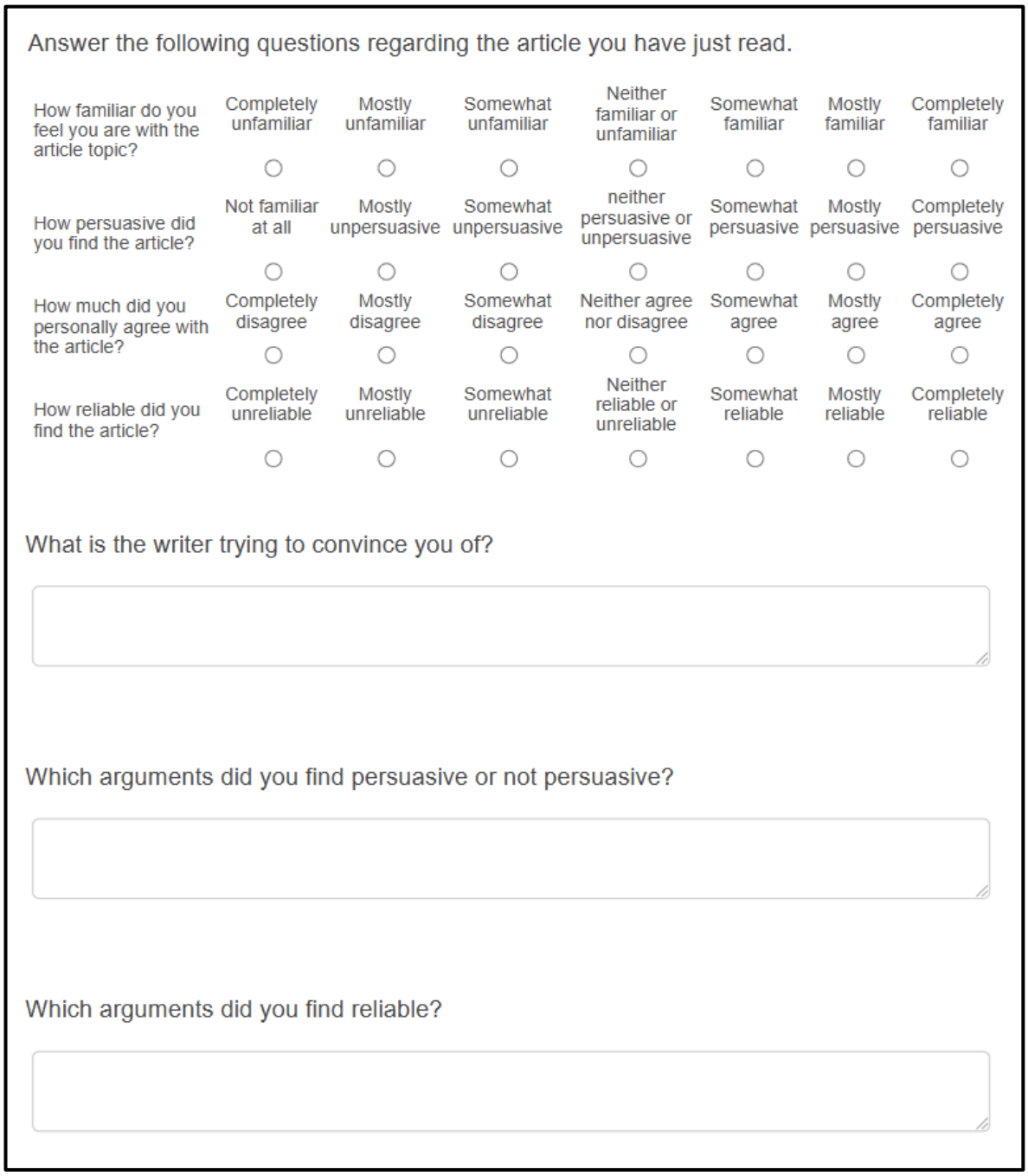
Priti Patel, the UK Home Secretary, has come under fire after introducing sudden tougher immigration regulation without warning. The new regulation will see the return of recently arrived refugees and asylum seekers back to their original countries. This is obviously devastating to all, but especially children, who will be turned around even if they have parents or other family currently in the United Kingdom.

The experts: "a question of morals"
Timothy Dunlop, an expert on immigration in Europe and the United Kingdom has called this decision completely immoral, explaining: "immigration routes to the United Kingdom from Europe and the wider world can take forever, whether it be by small boat, vehicle, or other. Many men, women, and children that have already packed up their belongings, risked their lives, and are currently en route will now be rejected."

Priti Patel has defended the decision, claiming the United Kingdom's security is at risk, however given her historical decisions on immigration and recent family issues one wonders if there lies an ulterior motive.

The public are outraged
Despite the backlash from human rights organisations in both the UK and Europe, the decision seems to now be set in stone, with the first arrivals already being turned around. The decision has received backing from the prime minister, signalling the secretive, inhumane direction intended by the government.

Appendix 8: Study 1 Survey



Answer the following questions regarding the article you have just read.

How familiar do you feel you are with the article topic?	Completely unfamiliar	Mostly unfamiliar	Somewhat unfamiliar	Neither familiar or unfamiliar	Somewhat familiar	Mostly familiar	Completely familiar
How persuasive did you find the article?	Not familiar at all	Mostly unpersuasive	Somewhat unpersuasive	neither persuasive or unpersuasive	Somewhat persuasive	Mostly persuasive	Completely persuasive
How much did you personally agree with the article?	Completely disagree	Mostly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Mostly agree	Completely agree
How reliable did you find the article?	Completely unreliable	Mostly unreliable	Somewhat unreliable	Neither reliable or unreliable	Somewhat reliable	Mostly reliable	Completely reliable

What is the writer trying to convince you of?

Which arguments did you find persuasive or not persuasive?

Which arguments did you find reliable?

Figure: Survey questions used in the first study of the thesis (see Chapters 4 and 5).

Appendix 9: Study 1 Interview Guide

Interview Structure

Interviews will be approximately 45 minutes in length, and will cover a number of key topics and areas. Some of these interview topics relate to original research questions, however many are new research questions after the first round of data collection was analysed. The key characteristics of the interviews are as follows:

- Interviews are approximately **45 minutes long**, however this may be shorter or longer at the discretion of the interviewer (interviews will not exceed 60 minutes).
- The interview will be run in a **semi-structured** fashion, and the interviewer will follow an *interview guide* rather than structured, fixed interview questions.
- Interviewees will be selected from the participants in the study to represent both a wide range of demographics, and participants with differing results. They are requested to interview but can decide not to take part if they wish.
- Interviewees will be paid **£12 Amazon Gift Vouchers** for their time, regardless of interview length.
- Interviewees will be asked which location they prefer to be interviewed in. Examples of the University of East Anglia, a coffee shop, or their own home will be given. They are welcome to suggest their own locations.
- Each interview will be recorded on an audio recording device and transcribed afterward.

Interview Topics

- Recognising elements of reflective learning occurring between the inoculation intervention and the longitudinal post-test.
- Understanding whether the testing procedure being supervised/non-supervised affects the participant judgements within testing sessions.
- The effects of choosing a within-subjects study design over a control-treatment design.
- How the number of players affected the intervention effectiveness.
- Effects of the use of a physical intervention as opposed to a digital one.

Interview Guide / Example Questions

Introducing Questions

- You played the game just over 9 weeks ago, what do you remember from the session?
- Why did you decide to take part in the study?
- What were your overall impressions from the session?

Direct Questions

- How do you feel your biases are influenced by your news consumption?
- How did playing the game affect you and your day-to-day actions, thoughts, and processes?
- Would you play this game again?
- How did you feel the number of players affected the session?

Follow-up Questions

- Get the interviewee to expand on their previous answers. Why? How? In what way?

Probing Questions

- Follow up through direct questioning. For example, "could you expand on this?", or "you said [this], what do you mean?"

Specifying Questions

- Very similar to probing questions but specifying detail. For example, "what did you do then?"

Figure: Interview guide for semi-structured interviews used in the first study of this thesis (see Chapter 5).

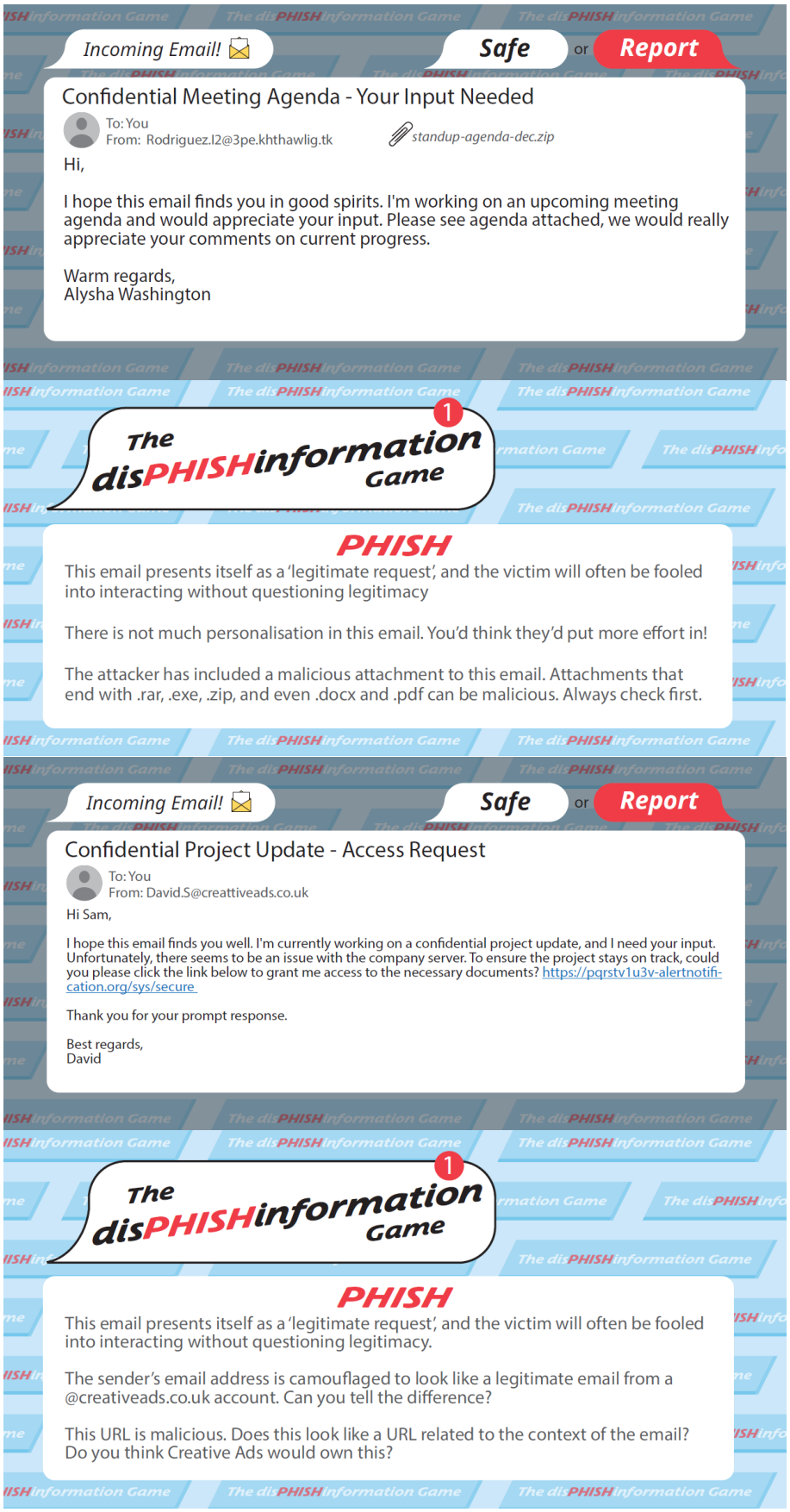
Appendix 10: Study 1 Example Coding Transcript

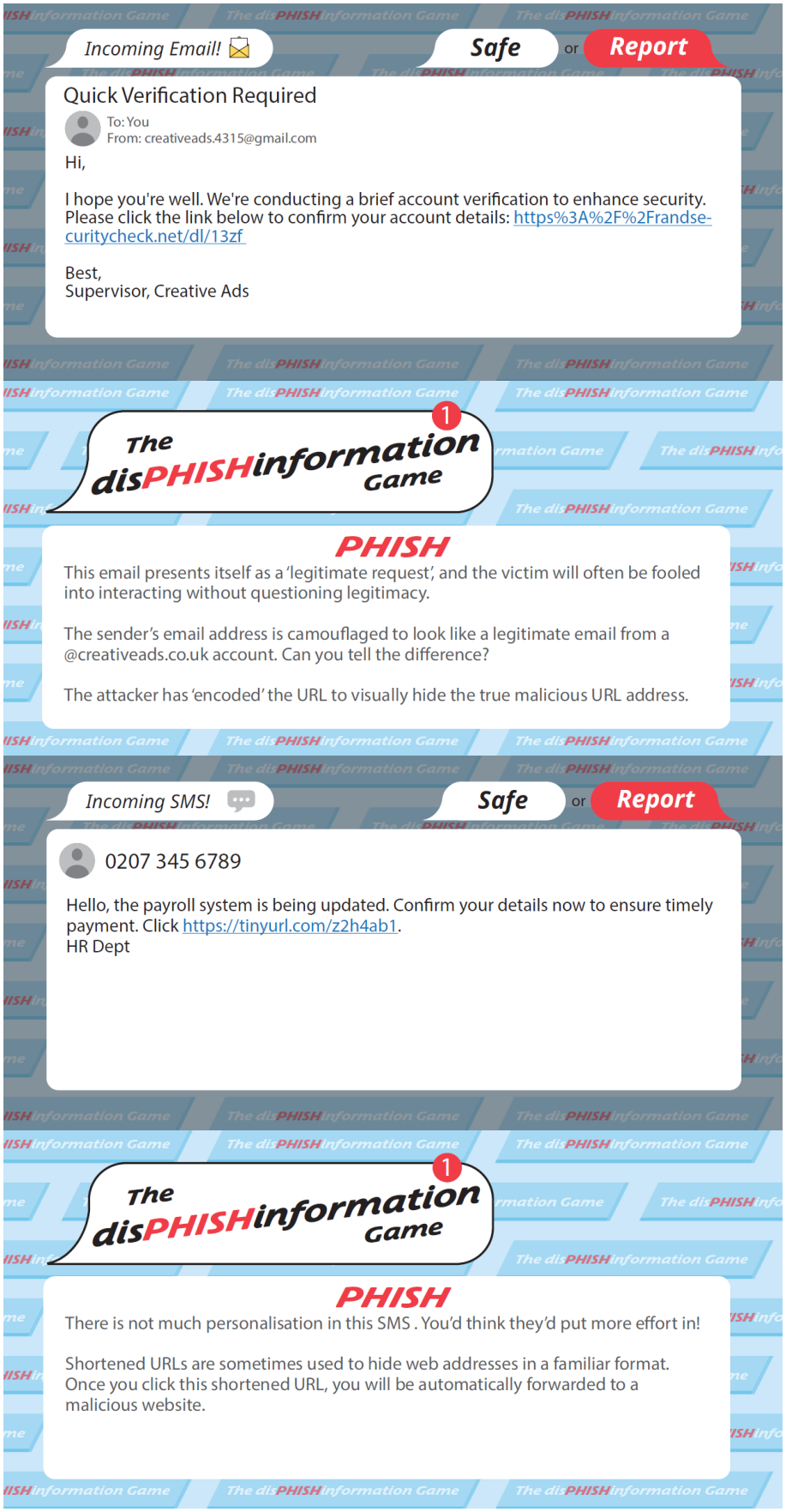
Table: Interview guide for semi-structured interviews used in the first study of this thesis (see Chapter 5).

00:05:54 Speaker 1 (researcher) Yeah, I see that. So now that I've sort of jogged your memory a little bit, reminded you of the session, I suppose one of my big questions is why did you decide to take part? 00:06:16 Speaker 2 (participant) I was curious really. 00:06:17 Speaker 1 (researcher) You can say for the money as well, that's alright. 00:06:20 Speaker 2 (participant) It was partly for the money, but mostly because I'm curious and I just wanted to see what it was about. because yeah, I like things like that. And and I really enjoyed it. Actually. I enjoyed the... working with people to work out what it was, and I found it interesting how like me and the other ladies my age, like totally agreed and we were like, yes, that's the answer. And then the older guys were... they were more on the same wavelength, and we disagreed with them sometimes. I found that really interesting, like where we were all coming from with that and maybe we were a bit more jaded, but yeah. 00:07:14 Speaker 1 (researcher) Oh, did you find it interesting the? Or do you remember discussing a lot with your teammates? 00:07:21 Speaker 2 (participant) I specific-, yeah, we talked the whole thing out all the way through. And there were some things- it was just really obvious- and then there was... there was one section I remember, I think the others just gave up and agreed with us because they didn't want to disagree any longer.	Codes used Participation motivation Group interaction discussion Note: 'group interaction discussion' code overlapping with 'participation motivation' code.
00:19:22 Speaker 2 (participant) Yeah, it was quite interesting the way I answered the first set of questions and then we played the game, and I was much more- I think I answered much more fully the second lot of questions than thought about things a lot more. And even though I know you can't necessarily- I already knew you couldn't necessarily trust an 'expert'. But it just made me really focus on that point a bit more and things like that. So yeah, it did make me think. Yeah, and it probably it probably has unconsciously carried on when I've read things afterwards. 00:20:15 Speaker 1 (researcher) Do you think so? 00:20:20 Speaker 2 (participant) Probably. Yeah, because. Yeah, I think about it, I think about the game quite a lot because that's just how I am Like I said, I don't really get out much. But it yeah- It's stuck with me. 00:20:35 Speaker 1 (researcher) It's interesting you saying earlier that you were actually [were] talking to a friend about the game. Do you mind if I ask what about?	**Codes used** Generic Identifying techniques Summarised experience Satisfaction or enjoyment Longitudinal effects

00:20:48 Speaker 2 (participant) I talked to people afterwards as well – like straight afterwards. 'Ohh I did this'; 'This was really interesting'; 'We did this and' – just because I just found it fascinating, and like I said, the dynamics between the four of us working it out together and how it worked. I can't think what we were talking about the other day. It just came up again and I was Like ohh yeah, I did this and Yeah, it was relevant.	Annotation: evidence of post-inoculation talk

Appendix 11: Materials from disPHISHinformation Game




Incoming Email!
Safe
or
Report
Quick Verification Required
To: You
From: creativeads.4315@gmail.com
Hi,
I hope you're well. We're conducting a brief account verification to enhance security. Please click the link below to confirm your account details: https%3A%2F%2Frandse-curitycheck.net/dl/13zf
Best,
Supervisor, Creative Ads
The disPHISHinformation Game
1
PHISH
This email presents itself as a 'legitimate request', and the victim will often be fooled into interacting without questioning legitimacy.
The sender's email address is camouflaged to look like a legitimate email from a @creativeads.co.uk account. Can you tell the difference?
The attacker has 'encoded' the URL to visually hide the true malicious URL address.
Incoming SMS!
Safe
or
Report
0207 345 6789
Hello, the payroll system is being updated. Confirm your details now to ensure timely payment. Click https://tinyurl.com/z2h4ab1.
HR Dept
The disPHISHinformation Game
1
PHISH
There is not much personalisation in this SMS . You'd think they'd put more effort in!
Shortened URLs are sometimes used to hide web addresses in a familiar format. Once you click this shortened URL, you will be automatically forwarded to a malicious website.

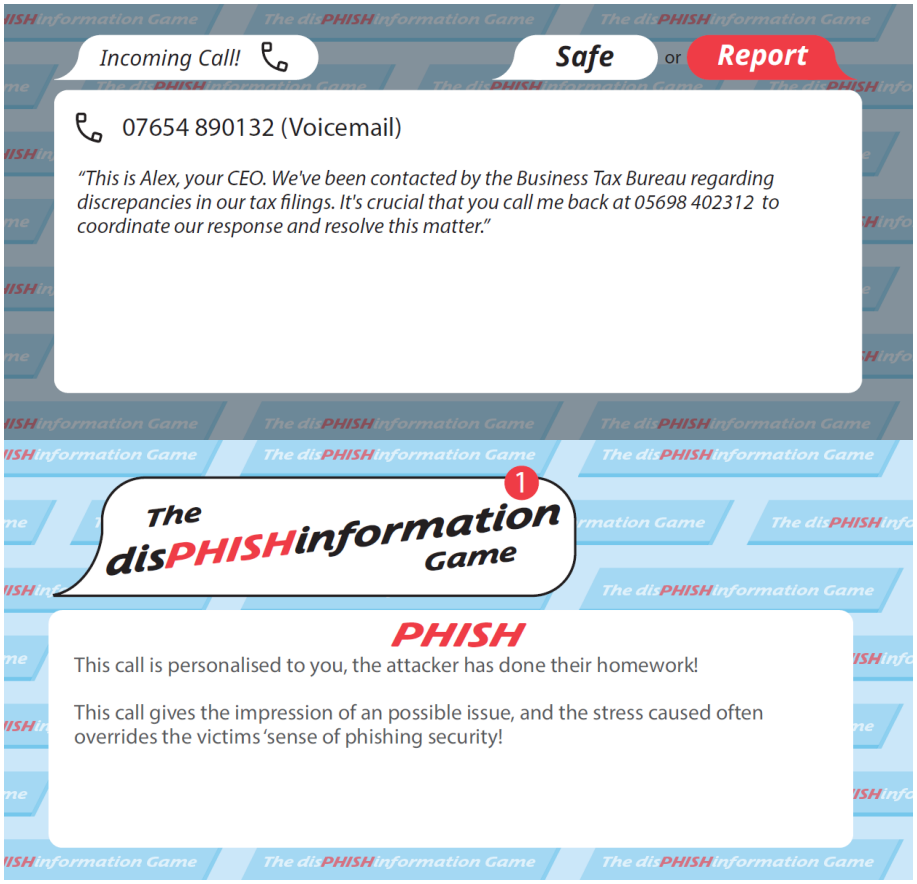


Figure: A sample of 'weakened dose' attack messages and corresponding refutations used in The disPHISHinformation Game.

Appendix 12: Session Observation Groups

Team number	Date	Group size	Character played	Interaction	Score
1	16/06/2023	2	Dennis in Denial	3.5	3
2	16/06/2023	2	Paula for Political Posturing	3	4
3	16/06/2023	2	Dana and Fox, Co-conspirators	4	2
4	16/06/2023	3	Mike the Money Changer	2	8
5	16/06/2023	2	Dennis in Denial	2.5	9
6	16/06/2023	3	Paula for Political Posturing	3	3
7	19/06/2023	2	Dana and Fox, Co-conspirators	2	7
8	19/06/2023	3	Mike the Money Changer	3	7
9	19/06/2023	3	Dennis in Denial	1	9
10	19/06/2023	2	Paula for Political Posturing	2	0
11	19/06/2023	4	Dana and Fox, Co-conspirators	3	2
12	20/06/2023	3	Mike the Money Changer	4	4
13	20/06/2023	3	Dennis in Denial	3	9
14	20/06/2023	3	Paula for Political Posturing	2.5	7
15	20/06/2023	2	Dana and Fox, Co-conspirators	3	3
16	20/06/2023	2	Mike the Money Changer	2.5	2
17	20/06/2023	2	Dennis in Denial	3.5	9
18	27/06/2023	4	Dennis in Denial	4	8
19	27/06/2023	3	Paula for Political Posturing	3	6
20	27/06/2023	3	Dana and Fox, Co-conspirators	3.5	9
21	27/06/2023	2	Mike the Money Changer	2.5	6
22	28/06/2023	3	Dennis in Denial	3.5	9
23	28/06/2023	2	Paula for Political Posturing	3	1
24	28/06/2023	3	Dana and Fox, Co-conspirators	2	4
25	28/06/2023	3	Mike the Money Changer	3	8