

TALLINN UNIVERSITY OF TECHNOLOGY
School of Information Technologies
Department of Software Science

Mauricio Antonio Duarte Lara 195459IVCM

PROTOTYPING A SERIOUS GAME ON INFORMATION MANIPULATION

Master's Thesis

Supervisors: María Claudia Solarte
Vásquez
PhD

Adrian Nicholas
Venables
PhD

Tallinn 2021

TALLINNA TEHNIKAÜLIKOOL
Infotehnoloogia teaduskond

Mauricio Antonio Duarte Lara 195459IVCM

**ÕPPE MÄNGU PROTOTÜÜBI LOOMINE
INFORMATSIOONI MANIPULEERIMISE
KOHTA**

Magistritöö

Juhendaja: María Claudia Solarte
Vásquez
PhD

Adrian Nicholas
Venables
PhD

Tallinn 2021

Author's declaration of originality

I hereby certify that I am the sole author of this thesis. All the used materials, references to the literature and the work of others have been referred to. This thesis has not been presented for examination anywhere else.

Author: Mauricio Antonio Duarte Lara

14.05.2021

Abstract

Adversarial actors leverage social media to achieve political objectives by employing information manipulation. This poses a risk to the confidentiality, integrity, and availability of information. These risks erode trust in institutions, distorts informed decisions, and affects democratic processes. A better defence can be obtained by a better understanding of adversaries. However, previous serious games on information manipulation have focused on psychological inoculation, ignoring the strategic motivations of adversaries in social media. With cybersecurity safeguarding information and operations in the context of adversaries, this thesis proposes to introduce policymakers to adversarial thinking through a Serious Game (SG). To achieve that aim, a document analysis was performed to identify the necessary considerations concerning learning, information manipulation, and serious games. The design of the prototype SG used a research design-oriented approach. The pilot testing employed an applied exploratory study with twelve participants, six from a legal background and the remaining six from an e-governance background. Data collection utilized two surveys with open-ended, multiple choice, and rating scale questions. Learning outcomes was measured by evaluating the participants' confidence levels on their definition of a concept. Given the sample size, the results are not conclusive. However, the data shows an increase in the confidence for all participants. This thesis has three key contributions. First, the application of the SG on the novel audience of policymakers. Second, the design of the prototype SG which incorporates the previously mentioned educational considerations. And last, further exploration on the contributions of cybersecurity to address information manipulation on social media.

This thesis is written in English and is 42 pages long, including six chapters, 19 figures and 8 tables.

Table of contents

Table of contents	5
List of figures	6
List of tables	7
Glossary	8
1 Introduction	9
2 Background.....	12
2.1 Social Media Information Manipulation and Cybersecurity	12
2.2 Learning Considerations for Serious Games	14
2.3 Serious Games	19
3 Methodology.....	25
4 Hashtag Struggle Design Process	26
4.1 Method.....	26
4.2 Prototype Design	26
4.3 Hashtag Struggle Overview	29
4.4 Discussion.....	36
5 Hashtag Struggle Pilot Test	38
5.1 Method.....	38
5.2 Experiment Design	39
5.3 Results	40
5.4 Findings and Discussion.....	45
6 Conclusions	51
Acknowledgements	53
References	54
Appendix 1	61
Appendix 2	63
Appendix 3	67
Appendix 4	69
Appendix 5	72

List of figures

Figure 1: Research question and contributions.....	10
Figure 2: The disinformation kill chain [42]	14
Figure 3: The Experiential Learning process, taken from [51]	16
Figure 4: Interplay of elements for learning in SGs [57]	17
Figure 5: The 1-2-1 model (top) and the Theoretically Informed Approach model (bottom) [60]	18
Figure 6: RQ1 and its research tasks.	26
Figure 7: Network graph, taken from[11]	28
Figure 8: Hashtag Struggle in Tabletop Simulator.....	29
Figure 9: An example of a distorted community on the left and a protected community on the right.....	31
Figure 10: Game board with communities and initial connections.	31
Figure 11: Possible connections between communities	32
Figure 12: Example of an engagement in #ELECTIONS. FWB attacks FNF. Each player loses one unit.	35
Figure 13: Hashtag Struggle at the start of the game	35
Figure 14: Sequence of play	36
Figure 15: Research question 2 and its research tasks	38
Figure 16: Piloting experiment outline.....	39
Figure 17 Pre-session versus post-session average confidence definition levels.....	47
Figure 18: Pre-session versus post-session average confidence definition levels by gender group.....	48
Figure 19: Pre-session versus post-session average confidence definition levels by age group.....	48

List of tables

Table 1: Current games on information manipulation.....	23
Table 2: Key Hashtag Struggle concepts.....	30
Table 3: Blue team actors, objectives, and units	32
Table 4: Red team actors, objectives, and units	33
Table 5: Player actions	33
Table 6: Pre-session adversarial thinking definitions verbatim.....	42
Table 7: Post-session adversarial thinking definitions verbatim	43
Table 8 : Post-session participants' main reflection verbatim.....	45

Glossary

Term	Definition
Active Measures	Covert or deceptive operations conducted in support of Soviet foreign policy to influence individuals, governments, or publics. [1]
Adversarial thinking	“The ability to embody the technological capabilities, the unconventional perspectives, and the strategic reasoning of hackers” [2]
Availability	“Ensuring timely and reliable access to and use of information.” [3]
Authenticity	“The property that data originated from its purported source” [4]
Confidentiality	“Preserving authorized restrictions on information access and disclosure, including means for protecting personal privacy and proprietary information” [3].
Cybersecurity	“A computing-based discipline involving technology, people, information, and processes to enable assured operations in the context of adversaries.” [5]
Disinformation	Intentionally misleading information [6].
Information	“Intentionally non-misleading representational content” [6].
Information pollution	“Irrelevant, redundant, unsolicited, and low-value information” [7].
Information laundering	Information laundering is legitimising false or deceitful information through the distortion it and obfuscation of the original source [8].
Integrity	“Guarding against improper information modification or destruction and includes ensuring information non-repudiation and authenticity” [3].
Infodemic	An overabundance of online and offline incorrect information to advance agendas of individuals or groups to undermine the public health response [9].
Malinformation	“Information that is based on reality, used to inflict harm on a person, organization, or country” [10].
Misinformation	Unintentionally misleading information [6].
Propaganda	Communication targeting a population to obtain a desired behaviour supporting the political goals of the propagandist [11].
Serious game	“Games with an explicit and carefully thought-out educational purpose and not intended to be played primarily for amusement” [12]
Social media	Online communication platforms that allow people to share different types of content to user-created networks [13].

1 Introduction

The information environment is disrupted by disinformation, misinformation, and malinformation. Digital wildfires [14], information pollution [7], and infodemic [9] are a few of the terms used to describe this disruption. From Ancient Egypt [15] to World War 2 Allied deception operations [16] to Cold War Soviet Active Measures [1], information has been mobilized in the pursuit of political objectives. This demonstrates that information manipulation is not a recent development, yet it is one taking new forms. Notably, the Internet Research Agency's efforts to shape public opinion during the 2016 United States election [17] pioneered social media information manipulation campaigns. Between 2017 and 2019, the Oxford Internet Institute has reported a 150% increase in countries conducting similar campaigns [18] underscoring the role of social media.

Social media is a prominent source of information. With a global penetration rate of 49% and 3.8 billion worldwide users [19], social media is a central platform for personal conversations, news updates, commercial communications, and political activism [20]. Social media refers to all online communication platforms used by people to share different types of content to user-created networks [13]. Some of its distinguishing features include its internet-based nature, its focus on content creation and sharing, and its network-enabled interactions [20]. Unlike traditional media, users are active participants [21] functioning as curators of information [22]. The end result is a constant, fast-paced, and pervasive environment [23] which requires users to rely on shortcuts [24] to evaluate the high volumes of information.

Information manipulation in social media compromises each attribute of the CIA triad. Adversarial actors have deliberately released confidential information at key points during elections to cause reputational damage to political candidates [17], [25]. Such disclosures target confidentiality. Concerning integrity, it is endangered not through improper modification or destruction, but by compromises to authenticity. The use of placement, layering, and integration techniques legitimizes manipulated information by obfuscating its true origin [8]. These techniques threaten authenticity, and thus integrity.

Lastly, availability is disrupted by overloading users with manipulated information using multiple channels and sources, an approach described as a firehose of falsehood [26]. While information is technically accessible, social media users might not find and access it, undermining availability.

Some of the consequences of manipulated information include citizens engaging in insecure actions including riots [27], lynching [28], and dying due to misleading medical information [29]. The European Commission acknowledges manipulated information has eroded trust in institutions, hampered citizen’s ability to make informed decisions, and affected policy-making processes [30]. Facing the threat of constant information manipulation campaigns, the government of Taiwan identified learning adversarial tactics as a countermeasure [31].

This thesis believes understanding adversarial tactics can be achieved through a Serious Game (SG). However, current SGs on the topic focus exclusively on the psychological aspects of information manipulation and do not have policymakers as their intended audience [32], [33]. Furthermore, these SGs ignore the adversarial nature driving information manipulation. This thesis proposes that gaining a greater understanding of information manipulation adversaries could lead to more effective policymaking. Cybersecurity as a discipline can contribute as it is the discipline concerned with safeguarding operations and information in the context of adversaries [5].

Research Questions	<p>RQ1: How to design the prototype of a serious game on information manipulation in social media?</p> <p>RQ2: How to pilot an experiment to evaluate the SG prototype’s game and educational elements?</p>
Contributions	<p>The application of the SG for the novel audience of policymakers.</p> <p>The design of the prototype SG which incorporates educational considerations like andragogy, transfer of learning, debriefings, and deployment.</p> <p>An exploration on the contributions of cybersecurity to address information manipulation on social media.</p>

Figure 1: Research question and contributions

The thesis is organized in six sections. Section 2 presents the background which elaborates on the intersection between information manipulation and cybersecurity, learning considerations for SGs, and SGs. Section 3 gives an overview of the methodology. Section 4 is the overview of the prototype SG, including the method used for its development and discussion. Section 5 concerns the method, design of the piloting experiment, its results, and findings and discussions. Section 6 provides the conclusions of this research.

2 Background

2.1 Social Media Information Manipulation and Cybersecurity

Information manipulation is composed of three types of manipulated information, with information being “intentionally non-misleading content” [6]. Manipulated information is a subtype of information that can be further categorized into disinformation, misinformation, and malinformation. Disinformation is intentionally misleading information [6]. Misinformation is unintentionally misleading information [6]. Malinformation is repurposed information with the intention to deceive [10], it is usually confidential information disclosed to cause harm [7]. While popular, the term fake news is to be avoided when referring to the subject of information manipulation. This term not only fails to capture the complexity of the issue, but more importantly it has been politically co-opted to dismiss disagreeable coverage [34].

Related to information manipulation, propaganda is a contentious term due to its political connotations which hinder its definition and discussion [35]. Propaganda is the communication targeting a population to obtain a desired behaviour supporting the political goals of the propagandist [11]. The persuasion of the targeted population does not necessarily involve the use of false or misleading information [36]. Propaganda is different from misinformation as the former is a deliberate effort while the latter is unintentional. By contrast, the difference between disinformation and propaganda is more subtle. While both are intentional and might use trustworthy information, disinformation is always misleading. In short, propaganda is a different concept that is politically motivated.

Cybersecurity is “a computing-based discipline involving technology, people, information, and processes to enable assured operations in the context of adversaries” [5]. While it is a computing-based discipline, cybersecurity incorporates interdisciplinary technical, human, organizational, and societal knowledge. The integration of these knowledge areas is achieved through transversal concepts such as confidentiality, integrity, availability, risk, adversarial thinking, and systems thinking [5]. Besides compromising the CIA triad, information manipulation on social media has other intersections with cybersecurity.

Social engineering is a type of attack, optionally supported by technical means, that exploits cognitive vulnerabilities using social interaction to achieve an objective [37]. To achieve the objective, social engineering attacks manipulate targets in performing actions that might not be in their best interest [38]. Reciprocation, commitment and consistency, social proof, liking, authority, and scarcity are the manipulation principles [39] that enable adversaries the exploitation of psychological vulnerabilities.

Information manipulation can be considered a social engineering technique because it operates in a similar manner. Social media provides adversaries with an information system with a global audience yet allowing the precision to reach users with a similar profile. Social media enables this precision due to their business model that relies on gathering data on its users. This business model provides the means for adversaries to not only tailor their messages, but to ensure they reach the intended audience. Lastly, information manipulation also exploits cognitive vulnerabilities to manipulate users in engaging in insecure behaviour.

In cybersecurity, adversarial thinking is embodying cyber adversaries by understanding their strategic motivations, technological capabilities, and unconventional perspectives [2]. The objective of adopting an adversarial thinking mindset is to improve defensive measures by gaining a better awareness on how adversaries operate. The cyber kill chain is a practical application of adversarial thinking. The cyber kill chain conceptualizes attacks as a series of sequential stages that adversaries execute to reach an objective [40]. The cyber kill chain helps defenders understand the different phases of an attack and the protective measures that can be applied during each. It reinforces the idea of placing adversaries at the centre when implementing strategic, tactical, and operational practices [40].

The cyber kill chain and similar cybersecurity concepts can be applied to disinformation campaigns due to the similarities between offensive cyber operations and information manipulation campaigns [41]. Like the cyber kill chain, conceptualizing information manipulation campaigns contributes to a better understanding of the attack and the available protective measures at each stage. The result would be a disinformation kill chain as proposed by [42], see Figure 2.



Figure 2: The disinformation kill chain [42]

Another cybersecurity concept with similarities in information manipulation concerns the categorization of cyber adversaries. Advanced Persistent Threats (APT) refers to sophisticated, well-resourced adversaries with political motivations behind their cyber operations. Similarly, an Advanced Persistent Manipulator (APM) [43] conceptualizes those adversaries conducting deliberate information manipulation campaigns to accomplish a politically motivated objective.

Two cases of the intersection between information manipulation and cybersecurity are the Doubleswitch case [44] and the 'Ghostwriter' influence campaign [45]. Concerning the former, Access Now reported how an adversary hacked social media accounts of prominent activists. Once compromised, the adversary took advantage of the reputation and influence associated with the original account to distribute manipulated information. As for the 'Ghostwriter' influence campaign, malicious actors compromised legitimate news websites to publish manipulated information. In some instances, this manipulated information was then referenced in social media by suspected fake personas. In both cases, adversarial actors conducted traditional cybersecurity attacks to compromise valuable targets who were then leveraged to support manipulated information campaigns on social media.

2.2 Learning Considerations for Serious Games

Learning is an enduring change in behaviour or in the capacity achieved through a form of experience [46]. Since learning is inferential, learning cannot be observed directly but

assessments can be used to determine the learning achieved by students. In addition, while learning is an enduring change it is not permanent as the learner can forget. There are several learning theories which differ on issues like the role of memory, motivation, learning process, transfer, self-regulation, and instruction implications [46]. However, learning theories share some common principles concerning the learner's progress, didactic material organization, feedback, the role of practice, and motivational factors. In learning, there are three main theories behaviourism, cognitivism, and constructivism [47], [48].

With behaviourism, learning is “[...] the acquisition of new behaviour” [49]. Under this theory, learning is focused on behaviours because they are observable learning outcomes [47] with mental processes discounted due to their unobservable nature. Behaviourism emphasizes the elements of the learning environment because they, and not the student, determine what can be learned [50]. Conditioning is the main learning method in this learning theory; mistakes are to be avoided as they are not considered a learning experience [47]. Learning is a gradual progress which must advance from simple to complex tasks.

Whereas behaviourism ignores cognitive processes, cognitivism puts them at the forefront of learning. Learning is an information handling and organization process where memory, knowledge, and representation are key aspects [47]. Cognitivism assumes learners have a schema, which is an existing knowledge representation students use for learning when processing new information [50]. Dissimilar from behaviourism, experimentation is part of the learning process by enabling insights the learner uses to arrive to a solution [47]. Like behaviourism, the learning process progresses from simple to complex. However, the role of reflexion is disregarded under cognitivism [47].

Constructivism shares the conception of learning as an internal process like cognitivism but differs on certain aspects. This learning theory conceptualizes learning as the construction of knowledge through experiences, with these constructions being unique to every learner [50]. Context and social interactions are relevant aspects in constructivism, with the former referred to as situated cognition. Situated cognition stipulates context must be considered when learning as it influences the process [47]. As for social interactions, conversations and shared problem solving between learners also support the construction of knowledge. Ideally, constructivist learning takes place in a context where

the learners' prior experiences will be leveraged and confronted with perspectives from other students [47].

Kolb introduced the Experiential Learning Theory (ELT) where learning is a four-stage cyclical progression transforming experiences into knowledge [51], see Figure 3. Learning in ELT is accomplished through the sequential completion of four learning modes: concrete experience, reflective observation, abstract conceptualization, and active experimentation. Learning can start in any of the modes if the cycle is completed. In ELT there are two opposite processes grouped in two distinct dimensions, with one dimension concerning the understanding of experience and the other its transformation. Abstract conceptualization and concrete experience are the two processes related to the understanding. Relatedly, reflective observation and active experimentation are the two processes covering the transformation of experience. For Kolb, learning is completed only when an understanding of experience is transformed [51]. On their own, neither understanding nor the transformation of experience are sufficient.

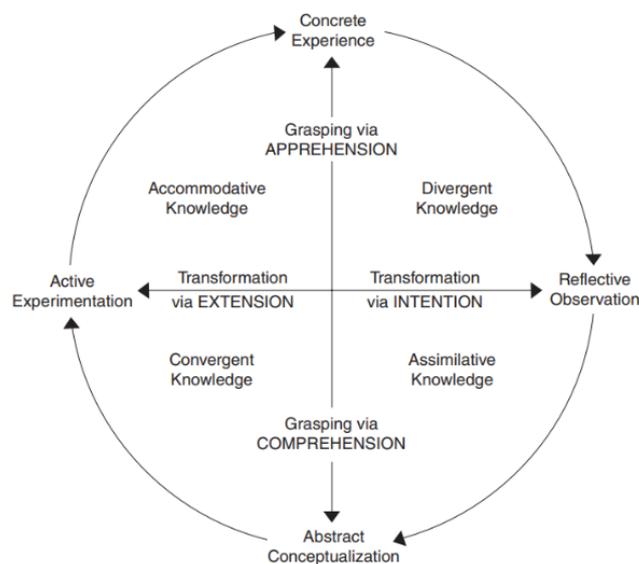


Figure 3: The Experiential Learning process, taken from [51]

Debriefing is essential for SGs because it provides space for reflection which is a requirement for completing the ELT cycle and thus, achieving learning [52]. Debriefing is a necessary step because it guarantees knowledge gains in the participants [53], [54]

regardless of the method used to conduct it [55]. Figure 4 shows Crookall's schema outlining the relationship between the elements of SGs involved in achieving learning. Finally, debriefing is needed to address the differences between the learning context and real-life [56] which is related to the concept of transfer.

$$[(Simulation/game + proper debriefing) \times engagement] = learning$$

Figure 4: Interplay of elements for learning in SGs [57]

The concept of transfer refers to the influence learning in one context has over a related performance in a different context [58]. Transfer of learning is relevant because education intends to create an effect beyond the classroom. Near and far transfer refer to the dimension of similarity between the learning context and the transfer context [46]. Near transfer implies a similarity between the original and transfer contexts, whereas in the case of far transfer the contexts are dissimilar. Low road transfer and high road transfer are two distinct mechanisms concerning the knowledge type. Low road transfer operates when the type of knowledge to be transferred is of a reflexive, spontaneous nature [46], [58]. By contrast, high road transfer applies to the types of knowledge that require a mindful and deliberate effort from the learner. Fast-paced interactions favour low road transfer while slower-paced activities are better suited for high road transfer [58]. Concerning SGs, transfer must be considered to ensure the game suits the educational objectives [59].

There are three models for deploying SGs and integrating their debriefings: the 1-2-1 model, the Theoretically Informed Approach (TIA), and the cloud model [60]. Of special interest are the first two models which will be focused on. The 1-2-1 model begins by a theoretical introduction, followed by a game session, and concludes with a debriefing and space for reflection, see Figure 5. The TIA approach starts in a similar fashion to the 1-2-1 model but differs by being cyclical. For each gaming session, there is a theory and reflection space with both repeating until the final reflection debriefing, see Figure 5. Thus, in terms of Kolb's learning cycle the 1-2-1 model allows one complete cycle iteration whereas the TIA allows for several iterations [60].

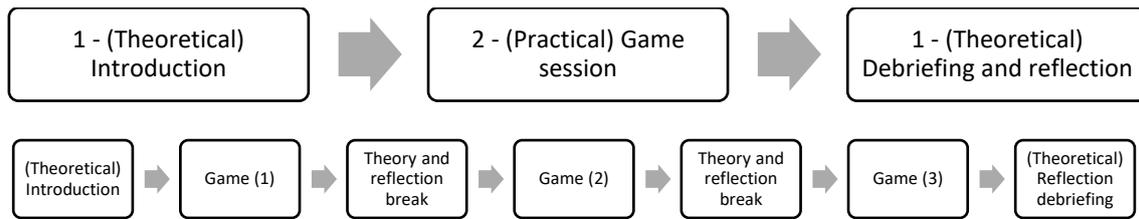


Figure 5: The 1-2-1 model (top) and the Theoretically Informed Approach model (bottom) [60]

Two additional considerations for learning involve the learner's characteristics, and the type of knowledge. Andragogy is the education focused on adults, with the assumption adults possess characteristics differentiating their learning from children's. Adults are understood as the individuals with socially productive roles, primarily responsible for their own lives, regardless of their age [61]. Introduced by Malcom Knowles, andragogy has four core assumptions to characterizing adult learners: self-direction, experience, readiness to learn, and learning orientation [61], [62]. Due to the responsibilities and self-direction required in adulthood, adult learners are self-directed learners who prefer to guide their education. Moreover, adults have acquired life experiences which are brought into the classroom. By connecting their experiences to new information, adult learners can be encouraged to adopt a more active role in the classroom [61], [63]. Adults want their education to be practical, with knowledge relevant to their current responsibilities [61]. In short, adult learners will value an active, relevant, and practical education which leverages their experiences.

As previously discussed, defining the type of knowledge to be taught is necessary to choose the most appropriate type of transfer. In Anderson and Krathwohl's revision of Bloom's taxonomy, educational objectives are split into knowledge and cognitive process dimensions [64]. The knowledge dimension describes the type of knowledge learners will acquire, while the cognitive process describes the process type. In this taxonomy, knowledge is classified into factual, conceptual, procedural, and metacognitive. As for the cognitive processes, they are categorized in an increasing order of abstraction and complexity: remember, understand, apply, analyse, evaluate, and create [65]. Educational objectives creation combines a cognitive process in the form of a verb with a knowledge type in the form of a noun.

2.3 Serious Games

Bernard Suits defined a game as the “[...] voluntary attempt to overcome unnecessary obstacles.” [66]. This definition can be expanded by including four defining characteristics present in every game: “[...] a goal, rules, a feedback system, and voluntary participation” [67]. The goal provides players with a purpose and guides their actions. The rules place limits on the actions available to achieve the goal, incentivizing creativity, and strategic thinking. To enjoy a better situation awareness, players require a mechanism to understand their current goal progress. The feedback system is the one responsible for providing this information through scores, points, or any other cues. Finally, voluntary participation refers to the players’ acceptance of the previous traits. This enables a common ground to play together. Moreover, the freedom to participate at will “[...] ensures that intentionally stressful and challenging work is experienced as safe and pleasurable activity.” [67].

Serious Games (SG) inherit all the previous four game characteristics but add a fifth one that differentiates them. SG are games with an explicit, though-out, educational purpose and with a primary objective other than entertainment [12]. Clark Abt introduced this concept in 1970 presenting a clear distinction of SG from other games with the inclusion of pedagogical elements [68]. The latter are the methods to achieve the learning objectives such as imparting knowledge, skills, or behaviour changes to the participants. Thus, SG “[...] aim to teaching something beyond the game play experience itself” [69].

Related to SG are the concepts of Game Based Learning (GBL), gamification, wargames, and simulations. GBL is the methodology of using games for educational purposes [70]. For some researchers SG and GBL are interchangeable concepts [71]. While GBL is a similar concept to SG, gamification is not. Gamification is the application of elements characteristic of games in non-game contexts [70] [72]. The goal of gamification is to induce individuals to engage in activities by making them more compelling. Gamification does not result itself in an end-product which further distinguishes it from SGs. Wargames are conceptualized representations of military conflict with a degree of authenticity regarding the combatants, their military capabilities, and combat related characteristics involved [73]. Since wargames have been used for training purposes, they overlap with SG depending on their educational content. Concerning simulations, they differ from SG on two aspects. Though SG might use some level of abstraction, simulations focus on

representing an aspect of real life with the highest possible fidelity. Furthermore, they do not incorporate any features related to education or entertainment.

The research on SGs as effective educational tools point towards positive findings. In a meta-analysis of SGs in education over a decade, the researchers found they have a positive impact on learning [74]. Another research not only supports SGs as effective learning tools, but adds they are better than traditional methods [75]. Certain characteristics of SG help explain their advantages. These can be summarized in the following three main arguments: engagement, educational sandbox, and social interaction.

SGs leverage play in addition to key structural elements like interaction, feedback, and goals to increase engagement [76]. Learning has been described as a challenging prospect, one that people tend to avoid in education or training [77]. Through play, SGs facilitate learning by putting players in a “[...] relaxed, receptive frame of mind for learning” [76]. The interaction found in SGs requires an active participation from the participants rather than the passive attitude more common in traditional education [78]. This interaction also means players receive feedback on their actions, sometimes eliciting emotionally arousing experiences that produce higher knowledge retention [79]. Finally, goals guide the efforts of players and provide a sense of purpose [67] which supports engagement.

SGs are an educational sandbox where participants explore the consequences of their actions in a safe environment. Certain contexts like crisis management, healthcare, or natural risk management are ill-suited for real-life experimentation due to their complexity, safety risks, or costs. In those contexts, SGs provide a safe environment where participants can practice and experience graceful failure in multi-patient care [80], natural disaster risk reduction [81], [82], or infectious disease control [83]. In other words, SGs allow participants to put into practice concepts that would remain theoretical notions only. Graceful failure [84] refers to the idea that failure is an expected, desirable, and even necessary outcome in the learning process [85], [86]. Thus, by removing the negative repercussions of failure, participants are encouraged to explore by taking risks and trying new approaches [87].

SGs are a conduit for social interaction, which can be leveraged for further educational purposes. Through cooperation or competition, SGs encourage exchanges between

participants resulting in learning moments [78]. Learning moments occur when the personal knowledge of the participants is prompted during the gaming session, leading to discussions and knowledge dissemination amongst the other players [78]. This dissemination goes beyond the gaming session, with participants exchanging information with their peers thus creating opportunities to inform hard to reach groups [88]. Furthermore, SGs can be used to provide a common language for the discussion of complex issues [89]. This common language can then be leveraged to integrate multidisciplinary knowledge and perspectives by integrating participants from a diversity of backgrounds [90].

In the field of cybersecurity, non-digital SGs have explored a diversity of subjects such as cyber conflict [78], cryptography [91], threat modelling [92], password security [93], social engineering [94], [95], industrial control systems security [96], and vulnerability exploitation [97] amongst others. The following showcase the state of the art concerning non-digital cybersecurity SGs.

Control-Alt-Hack is a card game with the educational goal of raising the awareness and altering perceptions on computer security [98]. The primary audience is undergraduate and engineering students in addition to high school students. Players assume the role of penetration testers competing to earn as many points as possible by completing missions. While the researchers report positive feedback from the educators using the game in the classroom, there is no material documenting the achievement of learning outcomes.

Decisions and Disruptions (D-D) is a cooperative SG where players manage the security of a cyber-physical environment [96]. Played in several rounds, players must reach a consensus on security acquisitions before facing pre-defined attacks chosen by the facilitator. The researchers created D-D with the purpose of exploring the security decision-making of different stakeholder groups. However, D-D also provided the players with a testing ground to experiment decision-making and reflect on their perceptions of security. From the gaming sessions, the researcher gained insights into how security experts make questionable decisions, the influence of individuals in decision-making and the dangers of assumptions.

[d0x3d!] is a SG with players acting as white-hat hackers infiltrating a network to retrieve digital assets [99]. To win, the players must cooperate to exfiltrate four digital assets while

remaining undetected. The adversary takes control of the fictional network, which reacts to the actions of the players, increasing the difficulty to accomplish the objectives. The goal of the designers is to expose K-12 students to computer security topics to increase their interest on the subject in an informal context. Nevertheless, the researchers acknowledge that the game's ability to meet its learning objectives has not been assessed properly.

Elevation of Privilege (EoP) is a card game with the goal of introducing threat modelling to developers [92]. EoP is based on the mechanics of Spades, with each card describing a threat belonging to one of the six categories of the STRIDE methodology. When a player plays a card, they must explain how the threat applies to the system that is being modelled. A scorekeeper not only tracks the points won, but also documents the threats to ensure the developers address the issues discovered through the game. Shostack aimed at using the structure provided by the game to teach developers threat modelling. Unfortunately, as for the examples cited above, no formal evaluation of the completion of learning outcomes is provided.

The Great (Cyber) Game is a wargame for cybersecurity education in the context of a cyber conflict between the United Kingdom and Russia [78]. Inspired by the UK National Cyber Security Strategy, players are split in two teams of three persons each. Each participant assumes the role of a distinct entity like the government, the electorate, or intelligence agencies amongst others. To win, each team must earn as many victory points as possible before the game ends. The educational goal is enabling learning moments amongst the participants for knowledge dissemination. The researcher provides several examples on how the game encouraged participants to discuss cybersecurity and its relationship to the wider context of society and politics.

Operation Digital Chameleon is a wargame where a red team develops an attack plan against a critical infrastructure protected by a Blue team [100]. Operation Digital Chameleon has a target audience of IT and IT security experts with the educational goal of exploring IT security in critical infrastructures. The game is played by two teams of three to six persons each and a facilitator who oversees the game. After drafting their plans, the facilitator reviews them with both teams present and determines a winner. Through the discussions facilitated by the debriefings, a new real-world attack vector was

conceptualized: CableJack. This discovery reiterates the exploratory qualities of SG that other researchers have pointed out.

Riskio is a tabletop SG designed with the purpose of increasing security awareness for non-technical audiences [101]. Hart et al. designed Riskio as an active learning environment for participants to practice cyber offense and defence in an easy to modify format. Their intended audiences are non-technical employees and first year cybersecurity university students. In turns, each player assumes the role of the attacker and formulates an attack based on a drawn card. The remaining players then must select an appropriate defence from their cards and describe how it counters the attack. Throughout the game, the facilitator provides comments, guidance, and ultimately decides which players chose the right defence.

In a summary, most non-digital SGs in this field are focused on traditional aspects of cybersecurity education, apart from the (Great) Cyber Game. In the games played with teams, teammates are motivated to work together by having a single, shared goal. Once more, the (Great) Cyber Game differentiates itself by having two teams where each player has different goals which creates a dynamic of tension inside the teams. However, this game presents a conflict with two clear opposing sides, ignoring the asymmetry present in cyberspace [59]. In addition, all previously mentioned games execute the briefing post-gameplay. The available options for deployment will be expanded in section 2.2. From the literature review, none of the games document considerations concerning the transfer of learning, nor cover the issue of information manipulation on social media.

Game Title	Medium	Players	Opponent
Bad News	Digital	1	Environment
Breaking Harmony Square	Digital	1	Environment
Go Viral!	Digital	1	Environment
The fake news game	Physical	1	Environment
UNISON	Physical	3 to 6	Environment

Table 1: Current games on information manipulation

On the topic of misinformation and disinformation several games have been developed, see Table 1. These include UNISON [102], The fake news game [32], Breaking Harmony

Square [33], Bad News [103], and Go Viral! [104]. A common characteristic found in the games is the focus on player versus environment dynamics. The players interact with an environment that responds to their actions but does not act as an adversary. Thus, there are no actors proactively challenging the players' efforts. Another commonality in four of the games is the goal of inoculating users from information manipulation through inoculation theory. None of the games explore the adversarial actors operating on social media, nor is their intended audience policymakers.

3 Methodology

The document analysis [105] conducted for the background included academic papers, reports, and books on the topics of information manipulation, cybersecurity, learning, and serious games. In addition to those sources, the document analysis also included non-academic documents like web articles and board games manuals. This analysis highlighted the need to use two complimentary yet distinct methods to answer the research questions. Both methods needed to support an iterative approach to gather feedback and identify improvements.

RQ1 required a design-oriented method to build the prototype SG. The experiences of other researchers supported the need for iterating, testing, and gathering feedback as part of the development process [82], [96], [102], [104]. Consequently, this thesis used a design science [106] approach. See section 4.1 for more details on the selected method.

RQ2 required a method to evaluate the prototype SG through an experimental approach. This thesis employed an applied descriptive study [107] to create a piloting experiment. This piloting experiment had the objective to evaluate the prototype and the experiment itself. Section 5.1 provides more detail on the method.

4 Hashtag Struggle Design Process

4.1 Method

RQ1 concerned the design of the SG prototype, see Figure 6. The objective was to produce a prototype that integrated the relevant theory considerations into its design. Thus, the method selected was a research design oriented approach [106], as used in the development of Operation Chameleon [100]. This approach intended to produce an artifact with a purpose. In this case, the artifact was the SG prototype which aimed to address a detected educational need. To support this process, this thesis applied techniques from design thinking. Design thinking supported prototyping for iterative evaluation and refinement [108]. This focus reinforced the need to create a prototype to be piloted in a prompt manner. This enabled gathering the necessary feedback for future iterations and research. Section 5 elaborates on the pilot testing of the prototype SG.

RQ1: How to design the prototype of a serious game on information manipulation in social media?	RT1: Design the SG prototype
	RT2: Create the SG prototype

Figure 6: RQ1 and its research tasks.

To adhere to the COVID19 safety guidelines, the SG switched from a physical board game to a digital one. The development of the prototype required the use of the software Tabletop Simulator to allow online play.

4.2 Prototype Design

The educational goal of the SG was to raise the awareness of policymakers on information manipulation in social media by teaching them adversarial thinking. Understanding how adversaries think is at the core of cybersecurity, and the first step towards better counter measures. A higher understanding by policymakers will lead to the formulation of more effective policies that address information manipulation in social media. It was assumed the intended audience had a low to moderate knowledge of information manipulation on social media. Thus, the SG had to avoid being too technical.

As the SG had to be developed during the process of a Master's thesis, board games offered advantages over digital games. The first concerns the simpler development offered by board games [78]. In board games, the gameplay is implemented by writing rules. This simplicity extends to the content creation and game modification. Should the need arise, incorporating a change can be accomplished while playing by editing the rules [78]. This is not possible in a digital game where similar modifications and testing are more complex and time-consuming activities.

Another advantage provided by board games is their inclusion of two elements benefiting learning and engagement: social interaction and the use of human as opponents. Constructivism proposed social interaction as beneficial to learning, see section 2.2. As the target audience were adults, andragogy also underscored social interaction as a favourable element for learning, see section 2.2. Concerning human opponents, their inclusion intended to add dynamism and uncertainty to the SG [78]. Both of those elements supported engagement.

After taking into consideration those advantages, the prototype benefitted from a board game as the format. With the game format settled, the next step was to define the type of knowledge to be taught. The category of knowledge determined the type of transfer, and thus the most suitable activity, see section 2.2. Using the revised Bloom taxonomy, it was classified as procedural, see section 2.2. This type of knowledge benefitted from high-road transfer which required slow-paced activities, see section 2.2.

With the type of activity identified, the next step was to outline a scope for the prototype SG. Establishing a scope was necessary to avoid making a highly detailed yet unplayable SG due to its complexity [59], [73]. Hamman and Hopkinson's definition of adversarial thinking [2] allowed a decomposition of information manipulation into three dimensions with different focuses. For this SG, the focus was the on strategic reasoning of adversaries. This thesis proposed strategic reasoning of information manipulation in social media involved focusing on the targeting of specific communities or topics by adversaries.

Communities were the contested space to highlight one of social media's most defining characteristics. Thus, communities became the board of the SG. However, the contested spaces had to be alterable to better represent the dynamic landscape of social media.

Network graphs were a source of inspiration for the aesthetics of the board, see Figure 7. Consequently, connections represented the interactions and communications formed between communities in social media. Players had to be able to alter them by establishing or removing connections between communities.

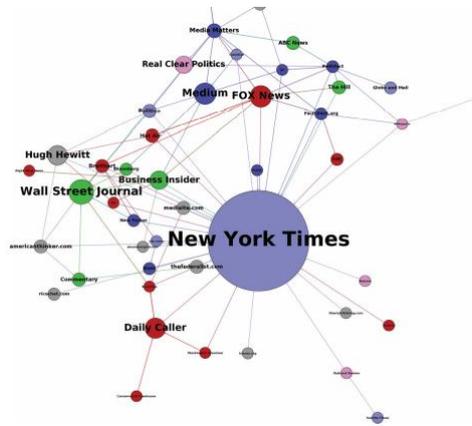


Figure 7: Network graph, taken from[11]

After defining the board of the game, the next step involved specifying the actors. Because the SG was adversarial, the actors belonged either to the Red or the Blue team. Nevertheless, while there were two teams, the objectives of each player had to be distinctive. This was inspired by the Great (Cyber) Game [78]. The reasoning was to add an element of internal conflict between team members. Unlike the previously mentioned game, the objectives were to be secret and disclosed only the intended player. The intention was to motivate players into anticipating their opponents' moves and determine their goals to reinforce adversarial thinking.

Concerning the Red team, each of the actors was meant to represent a different type of information manipulation adversary and their motivations. Benkler et al. [11] described politicians and fake news entrepreneurs as two actors with distinct causes partaking in information manipulation. The third team member was meant to represent an Advanced Persistent Manipulator (APM) [43] like the Internet Research Agency. The Red team had different communities to target as a representation of their different yet sometimes complementary financial, political, or strategic motivations.

As for the members of the Blue team, they included a social media platform as one of its members. This member had limited objectives assigned to convey their conflict of interest

when addressing information manipulation due to their business model. The East StratCom Task Force (ESCTF) inspired another team member to represent governments and other institutional actions to counter information manipulation. Lastly, Bellingcat motivated the design of the last member, which illustrated the debunking and fact-checking efforts of journalists and other non-governmental organizations.

Finally, the presence of an actor in a community meant to symbolize their efforts to protect or distort the information exchanged there. The Red team's presence had an advantage over the Blue team to underscore the difficulties in removing propagated manipulated information. Thus, the Red Team had to be able to distort a community even though they were not the majority. Attacking another player represented an abstraction on the use of botnets, trend hijacking, memes, and other techniques to propagate or remove information in a community.

4.3 Hashtag Struggle Overview

Hashtag Struggle Key Concepts

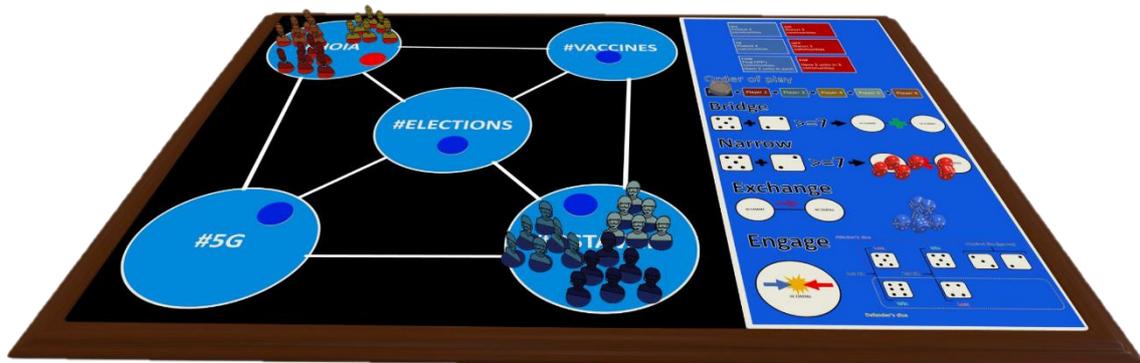


Figure 8: Hashtag Struggle in Tabletop Simulator

Hashtag Struggle is a contest over five social media communities in a fictional social media platform, see Figure 8. Six players, evenly split into the Blue and Red teams, clash to establish a presence in the communities. Each player controls a different social media information manipulation actor, each one striving to complete unique objectives. Appendix 1 presents a summary of the rules and a video extract from one of the playtesting sessions.

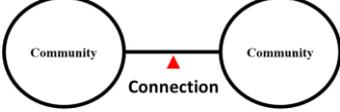
Concept	Definition	In-game representation
Unit	The pieces controlled by each player. Their presence in a community allows a player the execution of specific actions. Each player has a total of six units available throughout the game.	
Community	The terrain of Hashtag Struggle. There are five communities in total. Players must have units present in specific communities to achieve their goals.	
Connection	The bridge between two communities. It determines valid movement options for the players' units. Players can add or remove connections. Represented in-game by a white line.	

Table 2: Key Hashtag Struggle concepts

Table 2 outlines key Hashtag Struggle concepts. In Hashtag Struggle, the players belong to either the Blue or the Red team. The six players are meant to represent key participants related to information manipulation on social media. The team affiliation determines the players' goal with the Blue team protecting specific communities, while the Red team seeks to distort them. Regardless of the team, each player controls a total of six units. The units of the Blue team are shades of blue, whereas the Red team's ones are shades of red. Players use units to establish a presence in the communities. Establishing a presence in a community allows the addition or removal of connections and the elimination of opposing units present in the same community. In this prototype of the game, an additional participant is required to fulfil the role of the facilitator. The facilitator assigns each player a team, an information manipulation actor, their units, and objectives.

The number of a team's units present in a community determines its status with two mutually exclusive status for a community: protected or distorted. The status of a community is tracked in-game with a two-sided circular token. If the blue side is up, the community is protected. Otherwise, the red side means the community is distorted. A community is protected when the total number of blue units is greater than the total

¹ "soldier" by icon 54, used under CC BY / Recoloured from original.

¹ "anonymous" by icon 54, used under CC BY / Recoloured from original.

number of red units. Similarly, a community is distorted when the total number of red units is greater than or equal to the total number of blue units. Figure 9 shows an example of a protected community on the right. On the right, the community has four units from the Blue team versus three of the Red team. Consequently, its status is protected. A distorted community is shown on the left of Figure 9. The community is distorted because there are three red units present versus three blue units.

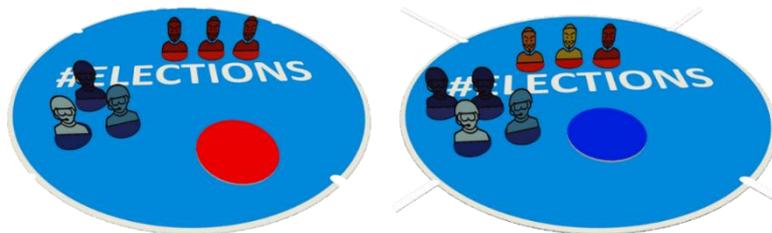


Figure 9: An example of a distorted community on the left and a protected community on the right

Figure 10 shows the game board at the start of the game with five social media communities and their default connections. Connections are bridges allowing player-controlled units to move between communities; they determine valid movement paths. During gameplay, players can add or remove connections between communities with certain restrictions.

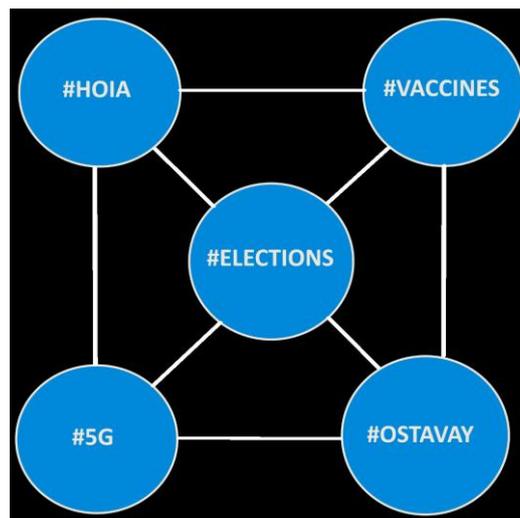


Figure 10: Game board with communities and initial connections.

First, a connection cannot be removed from a community if it is the last one. Second, there cannot be any duplicate connections. On the left of Figure 11 is an example of duplicate connections, with communities 1 and 2 connected through H and F which is invalid. On the right, connections A, B, and C can be removed from community 3 because it has four connections in total. However, connection D cannot be removed as it is the only connection of community 4.

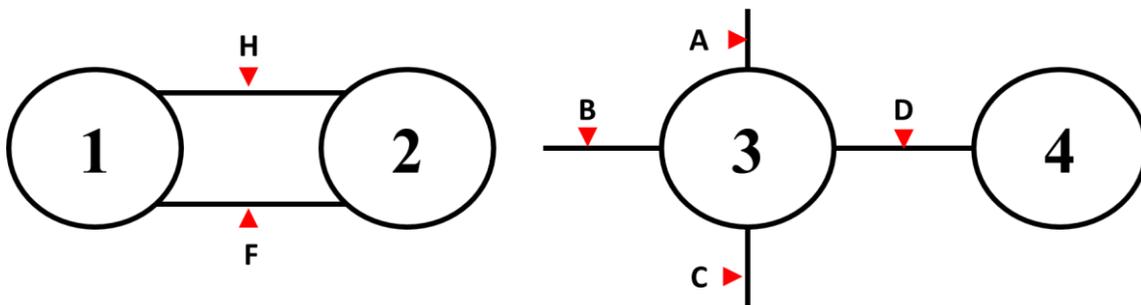


Figure 11: Possible connections between communities

Hashtag Struggle: Player Goals

Blue team members and objectives	Units
Baltic Information Agency (BIA). Objectives: Protect the following communities: #VACCINES, #ELECTIONS, and #5G.	
Factbook (FB). Objectives: Protect the following communities: #OSTAVAY and #VACCINES	
Facts Without Borders (FWB). Objectives: Find Hyper Partisan Party's communities. Have at least 2 units in each.	

Table 3: Blue team actors, objectives, and units

The Blue team actors include the Baltic Information Agency (BIA), Facts Without Borders (FWB), and Factbook (FB). BIA is inspired by the East StratCom Task Force from the European External Action Service. Their in-game goal is to protect the following communities #VACCINES, #ELECTIONS, and #5G. FB embodies a social media platform with the goal of protecting only two communities: #OSTAVAY and #VACCINES. Finally, FWB represents a non-governmental organization based on

Bellingcat, First Draft, and similar NGOs. Their objective is to protect the two communities that the Hyper Partisan Party player is attempting to distort. Table 3 summarizes the goals for the Blue Team.

Red team members and objectives	Units
Fake News Farmers (FNF). Objectives: Have 2 units in the following communities: #5G, #VACCINES, and #HOIA	
Hyper Partisan Party (HPP). Objectives: Distort the following communities: #HOIA and #OSTAVAY	
Agency of Internet Research (AIR). Objectives: Distort the following communities: #VACCINES, #ELECTIONS, and #HOIA.	

Table 4: Red team actors, objectives, and units

The Red Team actors include the Agency of Internet Research (AIR), Fake News Farmers (FNF), and Hyper Partisan Party (HPP). AIR represents an advanced persistent manipulator inspired by the IRA. The AIR player wins by distorting the #VACCINES, #ELECTIONS, and #HOAI communities. FNF are disseminators of misinformation with a financial motivation. To achieve their objective, FNF must have two units in three communities: #5G, #VACCINES, and #HOIA. Lastly, HPP represents a political party intent on provoking support through disinformation. Their objective is to distort both #HOIA and #OSTAVAY. Table 4 summarizes the goals for the Red Team.

Hashtag Struggle: Player actions

Action	Description
Bridge	Create a new connection between two communities. The player must have units present in the community from where the connection originates.
Narrow	Remove an existing connection between two communities. The player must have units present in the community where the connection will be removed. It cannot be used to remove the last remaining connection.
Exchange	Move or restore any number of player-owned units. Units move between two connected communities. When restoring units, they must be placed in communities with friendly units.
Engage	Engage allows the removal of opposing units in a specific community where the attacking player has units.

Table 5: Player actions

Table 5 shows the four main actions players can execute when it is their turn to play: bridge, narrow, engage and exchange. Bridge creates a new connection between any two communities. Narrow removes an existing connection. In both cases, the player must have at least one unit in the community from where the bridge or narrow action will be performed. To determine its outcome, the player rolls two six-sided dice with the action being successful if the result is greater than or equals to 7. As previously mentioned, the last connection of community cannot be removed. Exchange allows a player to move as many units as needed from one community to another if the communities are connected. Alternatively, exchange allows a player to restore units lost during an engagement which must be placed in communities where friendly units are present.

Engage allows the removal of units from a specific player of the opposing team. When engaging, the attacker states the community where the engagement is taking place and the player acting as the defender. Both the attacker and defender roll a six-sided dice for each unit present in the community where the engagement is occurring. The dices are paired by taking the highest value from each side until no more dice pairs can be formed. Within each dice pair the values are compared, with higher values defeating lower ones. For each defeat in a pair, the losing side must remove a unit from the community. In case of a draw the defender wins. Any unpaired dice are ignored. Figure 12 shows an example of an engagement, with FWB attacking FNF in #ELECTIONS. The FWB player rolls four die, one for each of his units on the community. The FNF player proceeds similarly, rolling a total of two dies. Two pairs are formed. In the first pair, FNF wins as his value is superior to the attacker's. In the second pair, FWB wins due to his value being greater than the defender's. The result of the engagement is both sides lost a unit each, leaving three remaining units standing for FWB and one for FNF.

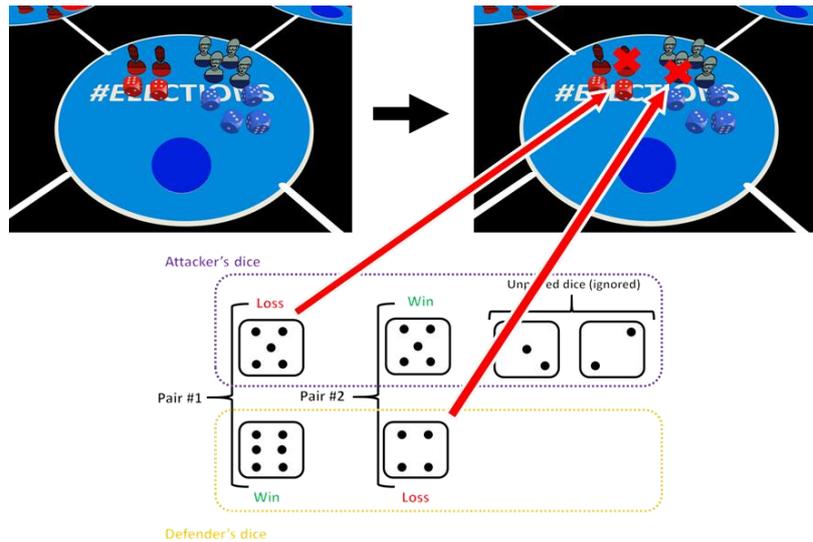


Figure 12: Example of an engagement in #ELECTIONS. FWB attacks FNF. Each player loses one unit.

Hashtag Struggle: Game start and sequence of play

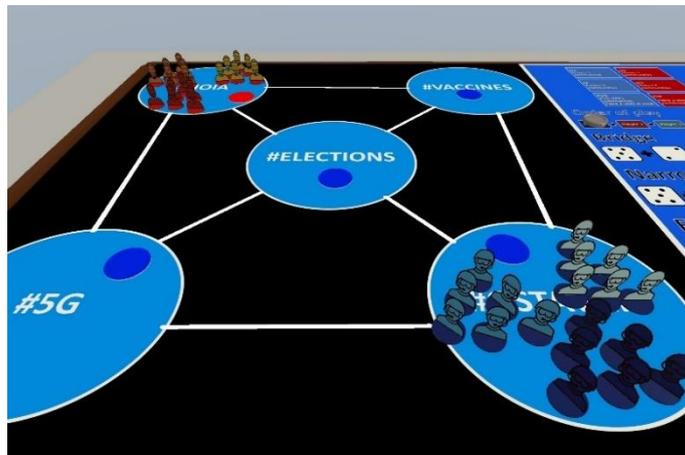


Figure 13: Hashtag Struggle at the start of the game

Figure 13 shows the start of Hashtag Struggle. During the setup of the game, the facilitator assigns each player an actor, their units, and objectives following the outline provided by Table 3 and Table 4. This designation should be done with the objectives of each actor revealed only to the designated player. All the units of the Blue team start in the #OSTAVAY community, whereas the Red team units begin in #HOIA. In addition, all communities except #HOIA should have their token status set to protected. Then, the

facilitator briefs the players on the key concepts of the game, the factions, and the sequence of play.



Figure 14: Sequence of play

Figure 14 shows the sequence of play, with each player consecutively executing a single action from the four available ones. A round ends when all six players have executed their actions. The suggested length of the game is 30 rounds.

4.4 Discussion

Transfer of learning is an educational aspect overlooked when developing SGs. The type of knowledge and transfer must be aligned, see section 2.2. From the research conducted, transfer of learning is not considered or has not been documented in the development of the reviewed cybersecurity SGs. In addition to being an important educational consideration, transfer of learning also supported the design of the prototype. By knowing high-road transfer required slower-paced activities, it was easier to consider or discard mechanics.

The main challenge on using a boardgame as the format concerned the balance between detail and abstraction. This challenge reflected the design trade-offs Shostack [92], Haggman [78] referred to when developing their games. To cover the Second Punic War, wargame designer Philip Sabin designed two different wargames[73]. One had a strategic focus, and the other a more tactical one. A similar approach could be applied to cover the different aspect of information manipulation in social media.

As for extending Hashtag Struggle, attribution is one extension. In the current iteration, the players' actions are easily attributable. This contrasts with the reality of cybersecurity and information manipulation where attribution is a challenge. One way to address this would be the use of social deduction and hidden role mechanics. In the Mafia/Werewolf

family type of games, players' actions are concealed. This leads players to make assumption on the responsible party behind an action. These mechanics could be leveraged to better represent the challenges of attribution.

Concerning information manipulation in social media, one concept to explore is the role of users. Golovchenko et al. research argues the average citizen is influential to both spreading manipulated information and countering it [22]. Citizens could be introduced in the form of another playable actor. Furthermore, this addition will help in challenging the traditional Red versus Blue contrast present in most cybersecurity SGs.

Another possible extension concerns the fog of war and the use of cards. Haggman used them in his wargame to add an element of imperfect information [78]. Cards could be incorporated in Hashtag Struggle to add imperfect information. In addition, the inclusions of cards could lead to discussions amongst the participants concerning, as reported by Haggman [78].

5 Hashtag Struggle Pilot Test

5.1 Method

RQ2 consisted in piloting an experiment to evaluate the SG prototype, see Figure 15. The objective of this evaluation was to analyse the educational and game aspects of the prototype. This analysis was needed to identify improvements for future iterations. RQ2 required a method that enabled the observation of the results of the prototype SG. For this reason, the most appropriate method was an applied exploratory study [107]. Specifically, an applied descriptive study as it allowed testing and evaluation of prototypes [107]. Another reason supporting this choice was the need to evaluate the piloting experiment and its data collection procedures. Applied descriptive studies supported this revision due to their relative setup ease [107]. The data collection involved the use of two online anonymous surveys.

RQ2: How to pilot an experiment to evaluate the SG prototype's game and educational elements?	RT6: Design the piloting experiment
	RT7: Conduct the piloting experiment
	RT8: Evaluate the SG prototype

Figure 15: Research question 2 and its research tasks

Sample Group

The sample group chosen for the experiment had to be available to participate in the experiment. Moreover, they had to fit the profile of future policy makers as they were the intended audience. No compensation or other incentives could be offered to the participants. Thus, the sample group consisted of volunteering students with legal and e-governance backgrounds. Twelve students in total participated, six from a legal background, and the remaining six from an e-governance background.

Data Collection

The data was collected using two online anonymous surveys. Both of the surveys had a mix of open-ended, multiple choice, and rating scale questions [107]. The use of surveys meant the collected data relied on self-reports by the participants. In addition, the use of open-ended questions presented a subjective opinion of the participants. To adhere to the

COVID19 pandemic safety guidelines, the piloting experiments had to be conducted online through Microsoft Teams. Finally, the sample size implied the results cannot be interpreted as conclusive.

5.2 Experiment Design

The experiment design aimed to incorporate debriefings and to evaluate the prototype SG. First, a deployment model must be defined to integrate the gaming sessions and debriefings. Learning is achieved when an ELT cycle is completed, see section 2.2. Thus, participants required a space for reflection which took the form of a debriefing. The TIA deployment model was chosen for two reasons. First, due to its novelty, see section 2.3. Second, the TIA allows the completion of more ELT cycles due to its cyclic nature, see section 2.2. With the deployment model chosen, the next step was to design the experiment, see Figure 16. The experiment's running time was one and a half hours.

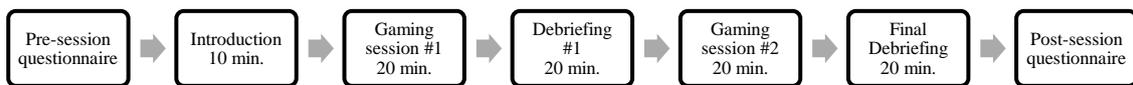


Figure 16: Piloting experiment outline

The experiment began with the completion of the first anonymous questionnaire by the participants, see Appendix 2. This questionnaire aimed to explore the awareness of participants on concepts related to information manipulation and cybersecurity. Question 8 and 9 intended to obtain a baseline on the learning outcomes before the participants played the SG. The type of knowledge to be imparted through the SG was procedural, see section 2.2. The assessment of procedural knowledge used the participant's level of confidence on their definition of a concept. Question 8 asked participants for a short definition of adversarial thinking. Question 9 requested the participants to rate their confidence on their previous answer being correct using a scale. The ranges of the scale went from one to five, with a value of one indicating no confidence, and five high confidence.

The facilitator explained the rules of the SG to the participants during the introduction, as well as answering any questions. The participants played the prototype, followed by the

mid-session debrief. The purpose of this debrief was threefold. First, the facilitator explained the concept of adversarial thinking and its relation to information manipulation in social media. Second, the facilitator addressed any questions the participants had. Third, to initiate a discussion with the participants.

With the first debriefing completed, the participants played the game one more time, followed by the final debrief. This debrief provided an additional space for reflection. The facilitator asked participants to describe their experience to further explore adversarial thinking and its relation to information manipulation in social media. In addition, the facilitator asked participants how the SG differs from reality. This difference needs to be addressed to avoid participants gaining false assumptions on the topic [54].

The final step was sending the second anonymous questionnaire, see Appendix 4. The second questionnaire intended to obtain feedback on the different components of the participants' experience. Questions 7 and 8 asked participants for their definition of adversarial thinking and to rate their confidence level on the answer. These questions measured the learning outcomes after the piloting experience. Question 9 requested the participants shared their main reflection on the experience. The goal was to provide an additional debrief in the form of a written response, see section 2.2.

5.3 Results

Pre-session questionnaire

75% of the participants identified as male, the other 25% identified as female. Concerning age, the split was 58.3% of participants having 25 to 34 years, and 41.7% between 18 to 24 years. 50% of the participants come from a legal background, the other 50% from an e-governance background.

Concerning the participant's education and training delivery, most participants received education or training through traditional methods with group discussions, e-learning, and instructor-led being the top three delivery methods. 33.3% of the participants were coached or mentored. Only one participant indicated GBL as an education or training delivery method. Same comment in here. One can only mention what stands out

Regarding the concepts related to information manipulation, all 12 participants specified being familiar with fake news. Misinformation came in second, with 91.7% of participants. 66.7% of the participants were familiar with disinformation. Malinformation is the concept participants were least familiar with only 25% of participants indicating an awareness.

When referring to intentionally misleading, false, or deceptive information, 58.3% of participants chose disinformation. 16.7 % chose fake news, 16.7% malinformation, and 8.3% misinformation.

Concerning unintentionally misleading, false, or deceptive information, 75% of participants chose misinformation, 8.3% disinformation, 8.3% malinformation, and 8.3% none of the previous concepts.

Confidentiality, integrity, availability, and risk are the concepts participants were most familiar with respectively 83.3%, 75%, 75%, and 58.3% of participants indicating familiarity with the concepts. 25% of the participants indicated being familiar with adversarial thinking, and 16.3% with system thinking. Only 8.3% indicated a total lack of knowledge of any of the previous cybersecurity concepts.

Participant	In 1 to 3 sentences, what do you think 'adversarial thinking' is about?
1	Putting yourself in the shoes of the person who's going to pen test your system for weaknesses/modelling the sort of attacks your system can expect to see
2	I assume it is linked with assuming someone's knowledge in some field. Not sure though, second time in my life that I hear this expression.
3	Hackers could maybe adapt the specific way of thinking of another, to copy one's strategy.
4	Knowing what to expect from cyber attackers & adversarial party
5	Thinking and thereby seeing things from another angle. This way helps to understand the motives of another person. For instance, a detective should try to see things the perspective of a person who committed a crime, this way it helps to identify the person, and what his motives could be.
6	The meaning of the concept is the capability to think for a step further like a hacker
7	Some action against somebody else, thought from the perspective in how this other would act

8	Thinking like the hacker
9	To think like the person we are performing an attack or we are securing from.
10	Trying to think in a similar way like a cyber criminal.
11	Adversarial thinking is a pattern of thinking that contradicts and challenges trend and narrative.

Table 6: Pre-session adversarial thinking definitions verbatim

Table 6 shows 11 participants submitted their definition of adversarial thinking. Of those, nine answers defined adversarial thinking in a similar fashion, describing it as a mindset of assuming the role of a hacker. Most of the participants related the concept with the field of cybersecurity except for one participant. This participant placed it in a broader context “[...] For instance, a detective should try to see things the perspective of a person who committed a crime [...]”. Of the remaining two participants, one acknowledged their lack of familiarity with the concept, defining it as assuming someone else’s knowledge. The other one responded with a definition unrelated to the others, specifying adversarial thinking as a mindset concerned with countering trends and narratives.

On the confidence levels the participants assigned to their definition of adversarial thinking, 33.3% of them indicated low confidence with their definition of adversarial thinking. A 25% showed an average level of confidence, and 16.7% a higher-than-average confidence. The remaining 25% indicated the highest level of confidence concerning their answer.

Post-session questionnaire

Relating to the fun element of the game, 75% of the participants strongly agreed the game was fun, with another 16.7% agreeing with the statement. However, 8.3% strongly disagreed with the statement and did not consider the game to be fun.

On the self-reported feeling of learning while playing, 16.7% of participants disagreed or strongly disagreed with the feeling of learning while playing. Another 16.7% were neutral concerning the statement. Lastly, 66.7% of participants agreed or strongly agreed with a feeling of learning while playing.

Concerning the participants' perception of the usefulness of the debriefings. 8.3% of participants disagreed on its usefulness, with another 16.7% being neutral. 8.3% agreed it was useful, and 66.7% strongly agreed with the briefing's usefulness.

About the areas of improvement, 58.3% chose an easier to understand game as the improvement. Another 41.7% considered a longer playing time as an improvement, 25% a more entertaining game, and another 25% a better integration of educational material. 8.3% considered the following improvements: better game components, more actions for the players, and a clearer end game.

Participant	In your own words, define adversarial thinking:
1	Thinking through the lens of someone whose interest might be to cause you harm.
2	Game theory.
3	You try to think like a cyber criminal and what a cyber criminal would do.
4	To think like a hacker, to understand what and how vulnerabilities can be attacked.
5	To think how the other will react in future decisions and act accordingly to have influence over the other person or community.
6	Basically trying to think about how other person might act\behave. I would not say that it's a "cybersecurity principle", I believe, it is a universal principle used in many different fields - criminology, psychology, as well as playing other games like chess, when you have to think in advance what might be the next step of the opponent.
7	What to expect from your adversary & from an attack.
8	The capability to comprehend atypical perspectives and tactical thinking of hackers
9	Thinking like the attacker/hacker who might want to mess with your system
10	"Thinking like a hacker". Basically it means thinking ahead and trying to predict the abilities of the other person.
11	Adversal thinking is about being ahead one step of others, continuously trying to anticipate the strategic or even big entities.

Table 7: Post-session adversarial thinking definitions verbatim

Table 7 presents the post-session participants' adversarial thinking definitions. 11 responses in total were received, with 5 definitions relating the concept to cybersecurity.

Of the remaining 6 definitions, 5 placed the concept in a broader context. Finally, one participant defined the concept as “*game theory*”.

About the participants’ confidence levels on their definitions, 8.3% of the participants did not feel confident on the correctness of their definition. 16.7% were neutral towards their confidence levels, 41.7% felt confident, and 33.3% felt very confident their definition was correct.

Participant	What is your main reflection from the whole experience?
1	Fun, Interesting and informative
2	Mauricio has the patience of a saint. The game is also legitimately fun, and would be more fun were it not for the constraints of academia
3	A great effort has been taken to plan the game. At first it was somewhat confusing what is the aim of the game, but later on it became more clearer. It was understandable what the game is trying to teach, but it would be better if there was more actions than move troops, cut or bridge lines and engage in a fight. Also, would be good if there was a definite ending point for the game.
4	That in terms of creating fake media, noise or volume makes a difference. The more people talking about it, the more it gets spread. Also the game taught us about how communities are interconnected with each other and hence that can make an impact in the communication outcome.
5	It was pretty fun, not sure the final purpose of the Master Thesis, but the game was fun.
6	The game is very fun, would be wise to finalize it and commercialize it:)
7	Part of the appeal was not to know what to expect. The strategical elements were present, point comes clear and overall the experience and communication was fun. The "game" elements might need some fine-tuning.
8	The game reminded me of the reality when the society needs to work in the form of a team, for instance, to work together for achieving the UN Sustainable Development Goals (in the game it was represented by the teams of red and blue, who were required to disinform and protect the fields respectively), however, each of the states involved in the common system has different objectives to reach for its own advantage (in the game each player had different objectives to achieve).
9	It was educational and interesting!
10	It was very interesting, fun and educational experience.
11	little bit confusing like in real life

12	Overall the piloting session experience was a top notch, something that would definitely be useful to combine into university studies. From the perspective of a lawyer, this kind of thinking is truly important and also overall in a fast developing world we need to adapt to new things, be ready to adapt new patterns in a multicultural environment. The educational lesson of the game also reached us, the players and personally I learned a lot new things. The element of fortune caused by the dices was also creative, as nothing in life can be straight forward calculated. All the best to you Mauricio and good luck with the thesis! :)
-----------	---

Table 8 : Post-session participants' main reflection verbatim

Table 8 presents the main reflection on the experience. A total of 12 responses were received. Of those, 9 responses focused on feedback related to either the game or the piloting session as the reflection. However, the responses from participants 4, 8, and 12 relate their piloting experience to aspects beyond the piloting experience. Participant 8 reflected on the similarities between the SG and reality, where the presence of different objectives in a team might lead to internal conflicts. Participant 4 pondered on the relevance of volume for the spread of manipulated information, and the influence of connections between social media communities. Lastly, participant 12 considered adversarial thinking to be relevant to her education as a lawyer because it enhanced her preparation.

In relation to the SG ratings, 8.3% rated the game with a 3, 16.7% gave it a score of 7, 33.3% a score of 8, 33.3% a score of 9, and 8.3% a score of 10. With regards to the session experience ratings from the participants. 8.3% rated the session with a 4, 16.7% gave it a score of 7, 16.7% a score of 8, 16.7% a score of 9, and 41.7% a score of 10.

Finally, 58.3% of participants indicated they would be strongly interested in further incorporation of SGs in their education or training, with another 25% indicating an above average interest. 8.3% are neutral concerning the proposition, and another 8.3% are not interested at all in SGs as an educational or training delivery method.

5.4 Findings and Discussion

Concerning the awareness on information manipulation, the pre-session questionnaire suggests the participants had an average knowledge on the topic. 91.7% and 66.7% of the

participants stated being familiar with misinformation and disinformation, respectively. However, while 91.7% indicated familiarity with misinformation, only 75% chose misinformation when referring to unintentionally misleading, false, or deceptive information. Similarly, 66.7% indicated being familiar with disinformation, yet only 58.3% chose disinformation when referring to intentionally misleading, false, or deceptive information. Furthermore, only 25% of participants were aware of malinformation.

The level of awareness varied depending on the demographic group. Concerning age, only 50% of the participants in the 18 to 24 years old group chose either disinformation or misinformation as the concepts being referred to. In the 25 to 34 years old group, 85.7% and 57.1% chose the right concept when referring to respectively, misinformation and disinformation. When considering educational background, the students with a legal background had a lower awareness level compared to the e-governance students. 50% and 66.6% of the former chose respectively disinformation and misinformation as the concepts being referred to. This is in contrast with the e-governance students, where 66.6% and 83.3% chose respectively disinformation and misinformation.

The previously mentioned findings hint participants are aware of information manipulation, but they lack clarity on the terms used to describe each of the different types of manipulated information.

Overall, participants had an awareness of basic cybersecurity concepts, with the results showing the CIA triad and risk were the most known concepts. As for adversarial thinking, only 25% of the participants indicated being familiar with the concept. Yet, the pre-session confidence level on the definitions had an average rating value of 3.3 out of 5. In addition, 41.6% of the participants felt confident or very confident concerning the correctness of their definition. The female participants had lower confidence levels than the male participants, with 2.3 being the average rating value for the former while the latter was 3.6. When divided by age group, the confidence levels were 2.8 for the 18 to 24 years old compared to 3.7 for the older age category.

Most of the participants defined adversarial thinking as a variation of thinking like a hacker. As defined in section 2.1, adversarial thinking includes three components: strategic motivations, technological capabilities, and unconventional perspectives. Most

of the definitions provided by the participants did not elaborate on its constituents, with some exceptions. For example, participant 5 mentions “[...] seeing things from another angle [...]” and “[...] understand the motives [...]” which can be related to unconventional perspectives and strategic motivations, respectively. Participant 7 included the perspective component in their definition “[...] thought from the perspective in how this other would act”.

The results from the post-session questionnaire suggest an increase in the confidence of the participants on the correctness of their adversarial thinking definitions, see Figure 17. Post-session, the average confidence rating was 4, compared to the 3.3 pre-session average.

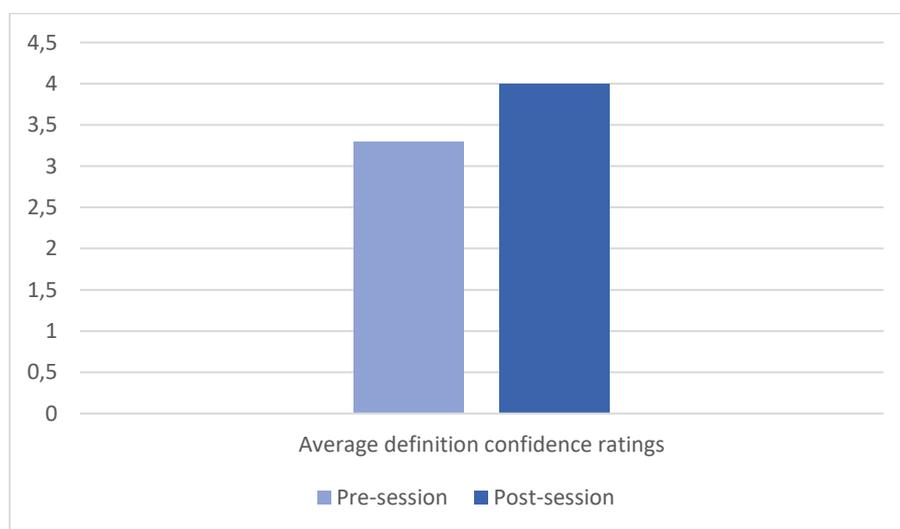


Figure 17 Pre-session versus post-session average confidence definition levels

This increase in confidence is found across the different demographic groups, see Figure 18. The average rating from female participants increased from 2.3 to 4, while male participants had theirs go from 3.6 to 4.

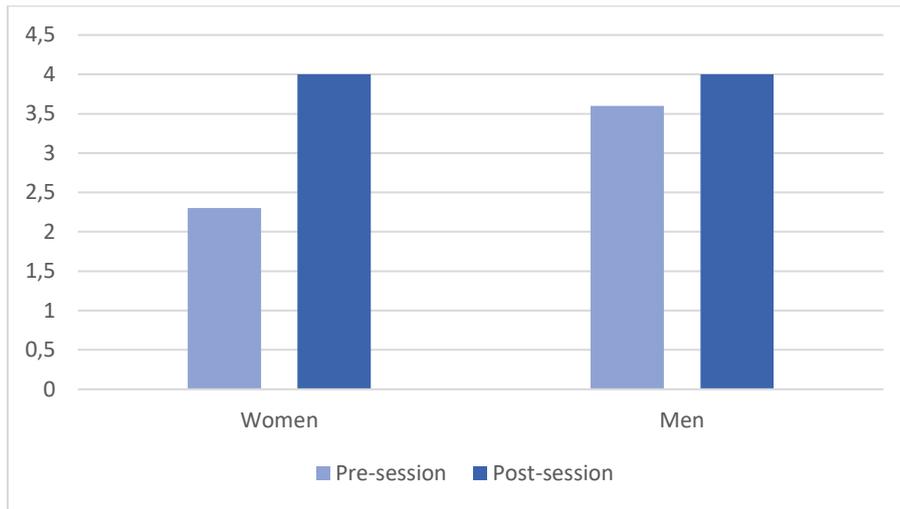


Figure 18: Pre-session versus post-session average confidence definition levels by gender group

The two age groups also had an increase, see Figure 19. The 18 to 24 group increasing from 2.8 to 3.8, while the 25 to 34 went from 3.7 to 4.14.

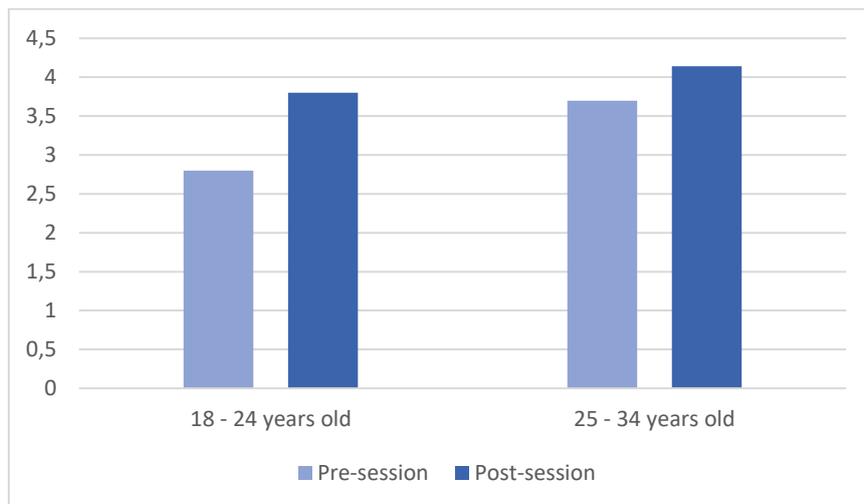


Figure 19: Pre-session versus post-session average confidence definition levels by age group

Unlike the confidence levels, the definitions of adversarial thinking did not change much. Like the pre-session definitions, most of the participants did not elaborate their post-session descriptions. Nevertheless, some of the definitions included elements either from the SG or the previously mentioned components of adversarial thinking. For example, participant 5 relates the concept to anticipating the actions to “[...] *have influence over*

the other person or community". Participant 8 integrates two components in their definition, with "*atypical perspectives*" and "*tactical thinking of hackers*" being included. A similar influence of the piloting session is exhibited in some of the reflections shared by the participants. Participant 4 mentions game elements such as communities and connections in their reflection. As for participant 12, the piloting experience related the different objectives present in the teams to the different interests present in efforts such as the UN sustainable development goals.

The experiment design used the participants' confidence levels on their definitions to measure procedural learning outcomes, see section 5.2. These results suggest the piloting experience might have achieved its intended learning outcome. However, these results are not conclusive due to the sample size and quantity of piloting experiments. More research will help in determining if confidence levels concerning a definition are a reliable method to measure of procedural types of knowledge.

In addition, the piloting experiment does not explore if this increased confidence was obtained by playing the SG, by the debriefings, or through a combination of both. The experiment could be changed to accommodate two variations. One, without debriefings, the other with. This comparison of results could help in confirming which component influences more the confidence levels and learning outcomes.

The difficulties participants had defining disinformation and misinformation highlights the challenges surrounding information manipulation. The first one concerns the lack of consensus on the umbrella term to contain disinformation, misinformation, and malinformation. Wardle and Derakhshan position those three concepts under the umbrella of information disorder [7]. However, others have used the terms digital wildfires [14], information manipulation [25], and infodemic [9] to refer to the same issue.

Furthermore, not only there is a lack of agreement concerning the umbrella term, but also on the definitions of the three types of manipulated information. This lack of consensus has been noted, with S e [6] and Baines and Elliot [10] proposing their definitions as a contribution to reach a consensus. Moreover, during the research for this thesis several researchers used the term fake news in their academic papers despite its contentiousness. If academics are having issues while agreeing on definitions, it is expected this will hamper non-experts when discussing information manipulation.

The participants' background has an influence on the areas of improvement requested for the SG. The developers of Riskio tested their SG with students and professionals [101]. The former group requested the addition of more game related elements, while the latter preferred an emphasis on the educational aspects. A similar experience occurred with Hashtag Struggle, with most of the participants, being students, requesting the improvement of game related elements. To address this issue, Hart et al. considered creating two different game boards, one for each audience [101]. These experiences suggest it is worthwhile to consider the participant's background before implementing their feedback. Thus, if the intended audiences for the SG are different, there might be a challenge in balancing the game and educational aspects.

Like transfer, debriefings and game deployment are two other aspects overlooked in cybersecurity SGs. The ELT cycle requires a space for reflection for the cycle to be completed which underscores the need for debriefings, see section 2.2. Of the reviewed cybersecurity SGs, only Rieb and Lechner documented using them [100]. However, their deployment model followed the traditional 1-2-1 model, with a single debriefing at the end of the session. Hashtag Struggle explored the use of the TIA model, which integrates debriefings not only at the end, but also between gameplay sessions.

6 Conclusions

This thesis proposed cybersecurity could contribute to address the issue of information manipulation in social media. Cybersecurity is the discipline of safeguarding operations and information in the context of adversaries. Information manipulation presents a risk to the confidentiality, integrity, and availability of information. In addition, there similarities between information manipulation and traditional cyber-adversaries. A better defence begins by a better understanding of adversaries and their modes of operations. A better understanding of their operations leads to better policymaking to address this issue.

This thesis aimed to design and pilot test a prototype SG on information manipulation in social media. A document analysis allowed the identification of key considerations concerning information manipulation, cybersecurity, learning, and gameplay components.

Research Question 1 concerned the design of a prototype SG on information manipulation in social media for policymakers. The intended goal of this prototype was to introduce participants to the cybersecurity concept of adversarial thinking. To ensure its educational objectives, the SG incorporated educational considerations like high road, social interaction, and active experimentation. In addition, the prototype leveraged hidden objectives, internal conflict, and human as opponents to engage the participants.

Research Question 2 concerned the design of the pilot testing of the prototype and its execution. The piloting experiment incorporated the Theory in Action model for the integration of debriefings in support to learning. The experiment used a pre-session and a post-session surveys to collect data. This experiment attempted to measure the procedural knowledge gain by measuring the participants' confidence levels on their definition of adversarial thinking. The surveys suggest an increase on their post-session confidence levels compared to their pre-session ones. Concerning the experience, the feedback from the participants was positive. The results indicated participants were engaged during the sessions, and they enjoyed the integration of SG and debriefings. The feedback also pointed out areas for further improvements concerning its educational and game aspects.

Designing any type of game is a challenging endeavour. The addition of educational elements adds another set of considerations which increase the complexity of the task. However, the experience provides an incorporation of synthesizing, creativity, and research skills like few others.

The design process underscored the relevance of educational aspects that are overlooked. Specifically, transfer of learning and debriefings are two educational elements which must be considered when designing an SG. These considerations ensure the educational outcomes are aligned with the game aspects.

Future work should explore in more detail the different methods to measure procedural knowledge learning outcomes. Another focus should be further explorations of applying the cybersecurity toolkit to the issue of information manipulation on social media.

Acknowledgements

This thesis was a personal journey, yet I did not accomplish it alone. Throughout this journey, several persons supported me. I would like to begin by expressing my gratitude to the European Union, the government of Estonia, TalTech, and all institutions who made it possible for me to study. I would also like to thank the cybersecurity master's admission committee for considering my application worthwhile of a scholarship.

Maria Claudia Solarte and Adrian Venables, I am grateful to you both for your professional, patient, and thorough guidance. A master thesis is a gruelling academic marathon. I am glad you supported me until the end. I would like to thank Peadar Callaghan from Tallinn University for his amazing one-hour master class on game design.

I would like to convey my eternal gratitude to my play testers for their willingness to participate and provide invaluable feedback. My fellow classmates, thank you: Sanjina Ershad, Mohamed Abdelmohsen, Risto Kasepu, Esteban Ramírez, and Burak Can Kus. To my favourite legal testers, thank so much: Niina Hakkinen, Antti Suomalainen, Vladislav Rusinov, Norman Aasma, Kyle Kimball, and Ivanna Tetera. Finally, special thanks to Serkan Koch, Marisu Cardenas and the e-governance crew: Norbert Ndashimye, Edi Kiviniemi, Thomas Ellert, Kakajan Kabayev. Cheers to Vasile Tarlev and Christman Roos for the interviews.

To my mom, dad, and sister which were my premium support network thank you for all the encouragements during these two years. It made the difference. Special mentions to la Señora Emma, Marie Christine, Adrián Lara, Víctor Duarte, Mileni Arraya, and Javier Alvarez. William Barrows, thank you for being my counsellor and being an endless source of support throughout these trials and tribulations.

To my friends, the journey would not have been the same without you: Sergio Trejos, Augusto Solís, Juan Carlos Cordero, Julián Gonzales, Antonio Alvarez, Erick Centeno, María José Chacón, Javier Zamora, Esteban Ramírez, Juan Manuel Delgado, Yoshihisa Furushita, Federico Martinazzi, Marco Jimenez, Adriana Umaña, Ekhterina Zhuchko, Carlos Cartín, Hector Madrigal, Christian Sánchez, Efraín Zeledon.

¡Pura vida!

References

- [1] T. Rid, *Active Measures: The Secret History of Disinformation and Political Warfare*. Farrar, Straus and Giroux, 2020.
- [2] S. T. Hamman and K. M. Hopkinson, “Teaching Adversarial Thinking for Cybersecurity,” *cisse*, vol. 4, no. 1, Art. no. 1, Oct. 2016.
- [3] *Public Law 113 - 283 - Federal Information Security Modernization Act of 2014*. Accessed: May 14, 2021. [Online]. Available: <https://www.congress.gov/113/plaws/publ283/PLAW-113publ283.pdf>
- [4] P. A. Grassi, M. E. Garcia, and J. L. Fenton, “Digital identity guidelines: revision 3,” National Institute of Standards and Technology, Gaithersburg, MD, NIST SP 800-63-3, Jun. 2017. doi: 10.6028/NIST.SP.800-63-3.
- [5] Joint Task Force on Cybersecurity E, *Cybersecurity Curricula 2017: Curriculum Guidelines for Post-Secondary Degree Programs in Cybersecurity*. New York, NY, USA: ACM, 2018. doi: 10.1145/3422808.
- [6] S. O. Søre, “Algorithmic detection of misinformation and disinformation: Gricean perspectives,” *Journal of Documentation*, vol. 74, no. 2, pp. 309–332, Jan. 2018, doi: 10.1108/JD-05-2017-0075.
- [7] C. Wardle and H. Derakhshan, “Information disorder: Toward an interdisciplinary framework for research and policy making,” *Council of Europe report*, vol. 27, 2017.
- [8] B. Carrasco Rodríguez, “Information Laundering in the Nordic-Baltic Region,” NATO STRATCOM COE, Nov. 2020.
- [9] “Managing the COVID-19 infodemic: Promoting healthy behaviours and mitigating the harm from misinformation and disinformation.” <https://www.who.int/news/item/23-09-2020-managing-the-covid-19-infodemic-promoting-healthy-behaviours-and-mitigating-the-harm-from-misinformation-and-disinformation> (accessed Mar. 28, 2021).
- [10] D. Baines and R. J. Elliott, “Defining misinformation, disinformation and malinformation: An urgent need for clarity during the COVID-19 infodemic,” *Discussion Papers*, vol. 20, 2020.
- [11] Y. Benkler, R. Faris, and H. Roberts, *Network propaganda: manipulation, disinformation, and radicalization in American politics*. New York, NY: Oxford University Press, 2018.
- [12] C. C. Abt, *Serious games*. University press of America, 1987.
- [13] N. Erragcha and H. Babay, “Social Media, Marketing Practices, and Consumer Behavior,” in *Leveraging Consumer Behavior and Psychology in the Digital Economy*, IGI Global, 2020, pp. 27–45.
- [14] L. Howell, World Economic Forum, and Risk Response Network, *Global risks 2013*. Cologny/Geneva, Switzerland: World Economic Forum, 2013.
- [15] A. Loktionov, “Ramesses II, victor of Kadesh: a kindred spirit of Trump?,” *the Guardian*, Dec. 05, 2016. Accessed: Dec. 13, 2020. [Online]. Available: <http://www.theguardian.com/science/blog/2016/dec/05/ramesses-ii-victor-of-kadesh-a-kindred-spirit-of-trump>

- [16] D. J. Bacon, "Second World War Deception: Lessons Learned for Today's Joint Planner," Air University Press, 1998. Accessed: Feb. 12, 2021. [Online]. Available: <https://www.jstor.org/stable/resrep13726>
- [17] R. DiResta *et al.*, "The Tactics & Tropes of the Internet Research Agency," p. 101.
- [18] S. Bradshaw and P. N. Howard, "The Global Disinformation Order 2019 Global Inventory of Organised Social Media Manipulation," Oxford Internet Institute, 2019. Accessed: Apr. 27, 2021. [Online]. Available: <https://demtech.oii.ox.ac.uk/wp-content/uploads/sites/93/2019/09/CyberTroop-Report19.pdf>
- [19] "Digital 2020: 3.8 billion people use social media," *We Are Social*, Jan. 30, 2020. <https://wearesocial.com/blog/2020/01/digital-2020-3-8-billion-people-use-social-media> (accessed Dec. 13, 2020).
- [20] H. Twetman, M. Paramonova, and M. Hanley, "SOCIAL MEDIA MONITORING: A PRIMER Methods, tools, and applications for monitoring the social media space," NATO STRATCOM COE, Dec. 2020.
- [21] J. H. Kietzmann, K. Hermkens, I. P. McCarthy, and B. S. Silvestre, "Social media? Get serious! Understanding the functional building blocks of social media," *Business Horizons*, vol. 54, no. 3, pp. 241–251, May 2011, doi: 10.1016/j.bushor.2011.01.005.
- [22] Y. Golovchenko, M. Hartmann, and R. Adler-Nissen, "State, media and civil society in the information warfare over Ukraine: citizen curators of digital disinformation," *International Affairs*, vol. 94, no. 5, pp. 975–994, Sep. 2018, doi: 10.1093/ia/iyy148.
- [23] A. M. Schejter and N. Tirosh, "'Seek the meek, seek the just': Social media and social justice," *Telecommunications Policy*, vol. 39, no. 9, pp. 796–803, Oct. 2015, doi: 10.1016/j.telpol.2015.08.002.
- [24] M. J. Metzger, A. J. Flanagin, and R. B. Medders, "Social and Heuristic Approaches to Credibility Evaluation Online," *Journal of Communication*, vol. 60, no. 3, pp. 413–439, Sep. 2010, doi: 10.1111/j.1460-2466.2010.01488.x.
- [25] J.-B. J. Vilmer, A. Escorcía, M. Guillaume, and J. Herrera, "Information Manipulation: A Challenge for Our Democracies, report by the of the Ministry for Europe and Foreign Affairs and the Institute for Strategic Research (IRSEM) of the Ministry for the Armed Forces," Policy Planning Staff (CAPS) and Institute for Strategic Research (IRSEM), Paris, 2018. Accessed: Mar. 13, 2021. [Online]. Available: https://www.diplomatie.gouv.fr/IMG/pdf/information_manipulation_rvb_cle838736.pdf
- [26] C. Paul and M. Matthews, "The Russian 'Firehose of Falsehood' Propaganda Model: Why It Might Work and Options to Counter It," Jul. 2016, Accessed: Nov. 30, 2020. [Online]. Available: <https://www.rand.org/pubs/perspectives/PE198.html>
- [27] "This Small Town Rioted Because Of Fake News And Rumors About The Coronavirus," *BuzzFeed News*. Accessed: Apr. 28, 2021. [Online]. Available: <https://www.buzzfeednews.com/article/christopherm51/coronavirus-riots-social-media-ukraine>
- [28] A. Gowen, "As mob lynchings fueled by WhatsApp messages sweep India, authorities struggle to combat fake news," *Washington Post*. Accessed: Mar. 06, 2021. [Online]. Available: https://www.washingtonpost.com/world/asia_pacific/as-mob-lynchings-fueled-by-whatsapp-sweep-india-authorities-struggle-to-combat-fake-news/2018/07/02/683a1578-7bba-11e8-ac4e-421ef7165923_story.html

- [29] “700 dead in Iran after drinking toxic alcohol to ‘cure coronavirus,’” *The Independent*, Apr. 28, 2020. Accessed: Apr. 28, 2021. [Online]. Available: <https://www.independent.co.uk/news/world/middle-east/coronavirus-iran-deaths-toxic-methanol-alcohol-fake-news-rumours-a9487801.html>
- [30] European Commission, “Tackling online disinformation: a European Approach, Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions.” Apr. 26, 2018. Accessed: May 13, 2021. [Online]. Available: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52018DC0236>
- [31] J. Blanchette, S. Livingston, B. Glaser, and S. Kennedy, “Protecting democracy in an age of disinformation: lessons from Taiwan,” Center for Strategic & International Studies, Jan. 2021. Accessed: May 08, 2021. [Online]. Available: <https://www.csis.org/analysis/protecting-democracy-age-disinformation-lessons-taiwan>
- [32] J. Roozenbeek and S. van der Linden, “The fake news game: actively inoculating against the risk of misinformation,” *Journal of Risk Research*, vol. 22, no. 5, pp. 570–580, May 2019, doi: 10.1080/13669877.2018.1443491.
- [33] J. Roozenbeek and S. van der Linden, “Breaking Harmony Square: A game that ‘inoculates’ against political misinformation,” *HKS Misinfo Review*, Nov. 2020, doi: 10.37016/mr-2020-47.
- [34] European Commission and Content and Technology Directorate-General for Communication Networks, “A multi-dimensional approach to disinformation: report of the independent high level group on fake news and online disinformation.” 2018. Accessed: Mar. 27, 2021. [Online]. Available: <https://data.europa.eu/doi/10.2759/739290>
- [35] É. Brown, “Propaganda, Misinformation, and the Epistemic Value of Democracy,” *Critical Review*, vol. 30, no. 3–4, pp. 194–218, Oct. 2018, doi: 10.1080/08913811.2018.1575007.
- [36] K. Born and N. Edgington, “Analysis of philanthropic opportunities to mitigate the disinformation/propaganda problem,” Hewlett Foundation, 2017. Accessed: Mar. 28, 2021. [Online]. Available: <https://hewlett.org/wp-content/uploads/2017/11/Hewlett-Disinformation-Propaganda-Report.pdf>
- [37] Z. Wang, L. Sun, and H. Zhu, “Defining Social Engineering in Cybersecurity,” *IEEE Access*, vol. 8, pp. 85094–85115, 2020, doi: 10.1109/ACCESS.2020.2992807.
- [38] C. Hadnagy, *Social engineering: The art of human hacking*. John Wiley & Sons, 2010.
- [39] R. B. Cialdini, *Influence: Science and practice*, vol. 4. Pearson education Boston, MA, 2009.
- [40] M. Muckin and S. C. Fitch, “A Threat-Driven Approach to Cyber Security,” p. 45, 2019.
- [41] C. R. Walker, S.-J. Terp, P. C. Breuer, and C. L. Crooks, PhD, “Misinfosec,” in *Companion Proceedings of The 2019 World Wide Web Conference*, New York, NY, USA: Association for Computing Machinery, 2019, pp. 1026–1032. Accessed: Feb. 05, 2021. [Online]. Available: <https://doi.org/10.1145/3308560.3316742>
- [42] DHS Analyst Exchange Program, “Combatting Targeted Disinformation Campaigns,” Department of Homeland Security, 2019. [Online]. Available: https://www.dhs.gov/sites/default/files/publications/ia/ia_combatting-targeted-disinformation-campaigns.pdf
- [43] C. Watts, “Advanced Persistent Manipulators, Part One: The Threat to the Social Media Industry,” *Alliance For Securing Democracy*, Feb. 12, 2019.

- <https://securingdemocracy.gmfus.org/advanced-persistent-manipulators-part-one-the-threat-to-the-social-media-industry/> (accessed Apr. 04, 2021).
- [44] “The ‘Doubleswitch’ social media attack: a threat to advocates in Venezuela and worldwide,” *Access Now*, Jun. 09, 2017. <https://www.accessnow.org/doubleswitch-attack/> (accessed Feb. 01, 2021).
- [45] Mandiant Threat Intelligence, “‘Ghostwriter’ Influence Campaign: Unknown Actors Leverage Website Compromises and Fabricated Content to Push Narratives Aligned with Russian Security Interests,” FireEye, Jul. 2020. Accessed: Feb. 01, 2021. [Online]. Available: <https://www.fireeye.com/content/dam/fireeye-www/blog/pdfs/Ghostwriter-Influence-Campaign.pdf>
- [46] D. H. Schunk, *Learning theories: an educational perspective*, 6th ed. Boston: Pearson, 2012.
- [47] P. Bélanger, “Three Main Learning Theories,” in *Theories in Adult Learning and Education*, 1st ed., Verlag Barbara Budrich, 2011, pp. 17–34. doi: 10.2307/j.ctvbkjx77.6.
- [48] M. Stewart, “Understanding learning: theories and critique,” *University teaching in focus: A learning-centred approach*, pp. 3–20, 2012.
- [49] A. Pritchard, *Ways of learning: Learning theories for the classroom*. Routledge, 2017.
- [50] G. McLeod, “Learning theory and instructional design,” *Learning Matters*, vol. 2, no. 3, pp. 35–43, 2003.
- [51] D. A. Kolb, *Experiential learning: Experience as the source of learning and development*. FT press, 2014.
- [52] D. Crookall, “Serious Games, Debriefing, and Simulation/Gaming as a Discipline,” *Simulation & Gaming*, vol. 41, no. 6, pp. 898–920, Dec. 2010, doi: 10.1177/1046878110390784.
- [53] M. A. Shinnick, M. Woo, T. B. Horwich, and R. Steadman, “Debriefing: The Most Important Component in Simulation?,” *Clinical Simulation in Nursing*, vol. 7, no. 3, pp. e105–e111, May 2011, doi: 10.1016/j.ecns.2010.11.005.
- [54] A. Ertan and P. Callaghan, “Principles of Effective Cybersecurity Wargames,” *Infosecurity Magazine*, Dec. 21, 2020. <https://www.infosecurity-magazine.com:443/next-gen-infosec/principles-of-effective/> (accessed Mar. 24, 2021).
- [55] C. Dufrene and A. Young, “Successful debriefing — Best methods to achieve positive learning outcomes: A literature review,” *Nurse Education Today*, vol. 34, no. 3, pp. 372–376, Mar. 2014, doi: 10.1016/j.nedt.2013.06.026.
- [56] V. A. M. Peters and G. A. N. Vissers, “A Simple Classification Model for Debriefing Simulation Games,” *Simulation & Gaming*, vol. 35, no. 1, pp. 70–84, Mar. 2004, doi: 10.1177/1046878103253719.
- [57] D. Crookall, “Engaging (in) Gameplay and (in) Debriefing,” *Simulation & Gaming*, vol. 45, no. 4–5, pp. 416–427, Aug. 2014, doi: 10.1177/1046878114559879.
- [58] D. N. Perkins, G. Salomon, and P. Press, “Transfer Of Learning,” 1992.
- [59] A. Ertan and P. Callaghan, “Enhancing Cyber Wargames: The Crucial Role of Informed Games Design,” *Strife*, Jan. 11, 2021. <https://www.strifeblog.org/2021/01/11/enhancing-cyber-wargames-the-crucial-role-of-informed-games-design/> (accessed Mar. 24, 2021).
- [60] T. D. Henriksen and T. Lainema, “Three approaches to integrating learning games in business education,” *Transforming University Teaching into Learning via Simulations and Games*, p. 15, 2012.

- [61] S. P. Forrest and T. O. Peterson, "It's Called Andragogy," *AMLE*, vol. 5, no. 1, pp. 113–122, Mar. 2006, doi: 10.5465/amle.2006.20388390.
- [62] D. D. Pratt, "Andragogy after twenty-five years," *New directions for adult and continuing education*, vol. 57, no. 57, pp. 15–23, 1993.
- [63] J. Collins, "Education Techniques for Lifelong Learning," *RadioGraphics*, vol. 24, no. 5, pp. 1483–1489, Sep. 2004, doi: 10.1148/rg.245045020.
- [64] L. W. Anderson and B. S. Bloom, *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. Longman, 2001.
- [65] D. R. Krathwohl, "A Revision of Bloom's Taxonomy: An Overview," *Theory Into Practice*, vol. 41, no. 4, pp. 212–218, Nov. 2002, doi: 10.1207/s15430421tip4104_2.
- [66] B. Suits, *The Grasshopper-: Games, Life and Utopia*. Broadview Press, 2014.
- [67] J. McGonigal and I. OverDrive, *Reality Is Broken*. Place of publication not identified: Penguin Group US, 2011. Accessed: Mar. 23, 2021. [Online]. Available: <http://api.overdrive.com/v1/collections/v1L2BaQAAAJcBAAA1M/products/44205f90-d66a-4b2b-8483-8a0cdb6a8822>
- [68] T. Susi, M. Johannesson, and P. Backlund, "Serious games: An overview," 2007.
- [69] K. Mitgutsch, "Serious Learning in Serious Games," in *Serious Games and Edutainment Applications*, M. Ma, A. Oikonomou, and L. C. Jain, Eds. London: Springer, 2011, pp. 45–58. doi: 10.1007/978-1-4471-2161-9_4.
- [70] K. Becker, "What's the difference between gamification, serious games, educational games, and game-based learning?," *Academia Letters*, Accessed: Mar. 03, 2021. [Online]. Available: https://www.academia.edu/45044609/What_s_the_difference_between_gamification_serious_games_educational_games_and_game_based_learning
- [71] K. Corti, "Games-based Learning; a serious business application," *Informe de PixelLearning*, vol. 34, no. 6, pp. 1–20, 2006.
- [72] S. Deterding, D. Dixon, R. Khaled, and L. Nacke, "From game design elements to gamefulness: defining 'gamification,'" in *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments*, New York, NY, USA, Sep. 2011, pp. 9–15. doi: 10.1145/2181037.2181040.
- [73] P. Sabin, *Simulating war: Studying conflict through simulation games*. A&C Black, 2012.
- [74] Y. Zhonggen, "A Meta-Analysis of Use of Serious Games in Education over a Decade," *International Journal of Computer Games Technology*, vol. 2019, p. 4797032, Feb. 2019, doi: 10.1155/2019/4797032.
- [75] S. de Freitas, "Are Games Effective Learning Tools? A Review of Educational Games," *Journal of Educational Technology & Society*, vol. 21, no. 2, pp. 74–84, 2018.
- [76] M. Prensky, "Fun, play and games: What makes games engaging," *Digital game-based learning*, vol. 5, no. 1, pp. 5–31, 2001.
- [77] J. Breuer and G. Bente, "Why so serious? On the relation of serious games and learning," *Journal for Computer Game Culture*, vol. 4 (1), pp. 7–24, 2010.
- [78] A. Haggman, "Cyber Wargaming: Finding, Designing, and Playing Wargames for Cyber Security Education," PhD Thesis, Royal Holloway, University of London, 2019.
- [79] L. Chittaro and F. Buttussi, "Assessing Knowledge Retention of an Immersive Serious Game vs. a Traditional Education Method in Aviation Safety," *IEEE*

- Transactions on Visualization and Computer Graphics*, vol. 21, no. 4, pp. 529–538, Apr. 2015, doi: 10.1109/TVCG.2015.2391853.
- [80] D. Tsoy *et al.*, “Creating GridlockED: A Serious Game for Teaching About Multipatient Environments,” *Academic Medicine*, vol. 94, no. 1, pp. 66–70, Jan. 2019, doi: 10.1097/ACM.0000000000002340.
- [81] S. Mossoux, A. Delcamp, S. Poppe, C. Michellier, F. Canters, and M. Kervyn, “Hazagora: will you survive the next disaster? A serious game to raise awareness about geohazards and disaster risk reduction,” *Natural Hazards and Earth System Sciences*, vol. 16, no. 1, pp. 135–147, Jan. 2016, doi: 10.5194/nhess-16-135-2016.
- [82] F. Taillandier and C. Adam, “Games Ready to Use: A Serious Game for Teaching Natural Risk Management,” *Simulation & Gaming*, vol. 49, no. 4, pp. 441–470, Aug. 2018, doi: 10.1177/1046878118770217.
- [83] J. Smith, N. Sears, B. Taylor, and M. Johnson, “Serious games for serious crises: reflections from an infectious disease outbreak matrix game,” *Globalization and Health*, vol. 16, no. 1, p. 18, Mar. 2020, doi: 10.1186/s12992-020-00547-6.
- [84] J. L. Plass, B. D. Homer, and C. K. Kinzer, “Foundations of Game-Based Learning,” *Educational Psychologist*, vol. 50, no. 4, pp. 258–283, Oct. 2015, doi: 10.1080/00461520.2015.1122533.
- [85] M. Kapur, “Productive Failure,” *Cognition and Instruction*, vol. 26, no. 3, pp. 379–424, Jul. 2008, doi: 10.1080/07370000802212669.
- [86] M. Kapur and K. Bielaczyc, “Designing for Productive Failure,” *Journal of the Learning Sciences*, vol. 21, no. 1, pp. 45–83, Jan. 2012, doi: 10.1080/10508406.2011.591717.
- [87] B. Hoffman and L. Nadelson, “Motivational engagement and video gaming: a mixed methods study,” *Education Tech Research Dev*, vol. 58, no. 3, pp. 245–270, Jun. 2010, doi: 10.1007/s11423-009-9134-9.
- [88] R. H. Bosma *et al.*, “Changing opinion, knowledge, skill and behaviour of Vietnamese shrimp farmers by using serious board games,” *The Journal of Agricultural Education and Extension*, vol. 26, no. 2, pp. 203–221, Mar. 2020, doi: 10.1080/1389224X.2019.1671205.
- [89] D. Reckien and K. Eisenack, “Urban Sprawl: Using a Game to Sensitize Stakeholders to the Interdependencies Among Actors’ Preferences,” *Simulation & Gaming*, vol. 41, no. 2, pp. 260–277, Apr. 2010, doi: 10.1177/1046878108321871.
- [90] K. Eisenack, “A Climate Change Board Game for Interdisciplinary Communication and Education,” *Simulation & Gaming*, vol. 44, no. 2–3, pp. 328–348, Apr. 2013, doi: 10.1177/1046878112452639.
- [91] A. I. González-Tablas, M. I. González Vasco, I. Cascos, and Á. Planet Palomino, “Shuffle, Cut, and Learn: Crypto Go, a Card Game for Teaching Cryptography,” *Mathematics*, vol. 8, no. 11, Art. no. 11, Nov. 2020, doi: 10.3390/math8111993.
- [92] A. Shostack, “Elevation of privilege: Drawing developers into threat modeling,” 2014.
- [93] H. Tupsamudre *et al.*, “GAP: A Game for Improving Awareness About Passwords,” in *Serious Games*, Cham, 2018, pp. 66–78. doi: 10.1007/978-3-030-02762-9_8.
- [94] K. Beckers and S. Pape, “A Serious Game for Eliciting Social Engineering Security Requirements,” in *2016 IEEE 24th International Requirements Engineering Conference (RE)*, Sep. 2016, pp. 16–25. doi: 10.1109/RE.2016.39.

- [95] D. Aladawy, K. Beckers, and S. Pape, “PERSUADED: Fighting Social Engineering Attacks with a Serious Game,” in *Trust, Privacy and Security in Digital Business*, Cham, 2018, pp. 103–118.
- [96] S. Frey, A. Rashid, P. Anthonysamy, M. Pinto-Albuquerque, and S. A. Naqvi, “The Good, the Bad and the Ugly: A Study of Security Decisions in a Cyber-Physical Systems Game,” *IEEE Transactions on Software Engineering*, vol. 45, no. 5, pp. 521–536, May 2019, doi: 10.1109/TSE.2017.2782813.
- [97] H. Enriquez, Y. Kadobayashi, and D. Fall, “Project config. play a turn-based strategy security board game,” in *Proceedings of the 12th European conference on games based learning (ECGBL 2018)*, 2018, pp. 72–81.
- [98] T. Denning, A. Lerner, A. Shostack, and T. Kohno, “Control-Alt-Hack: the design and evaluation of a card game for computer security awareness and education,” in *Proceedings of the 2013 ACM SIGSAC conference on Computer & communications security*, New York, NY, USA, Nov. 2013, pp. 915–928. doi: 10.1145/2508859.2516753.
- [99] M. Gondree and Z. N. J. Peterson, “Valuing Security by Getting [d0x3d!],” p. 8.
- [100] A. Rieb and U. Lechner, “Operation Digital Chameleon: Towards an Open Cybersecurity Method,” in *Proceedings of the 12th International Symposium on Open Collaboration*, New York, NY, USA, Aug. 2016, pp. 1–10. doi: 10.1145/2957792.2957800.
- [101] S. Hart, A. Margheri, F. Paci, and V. Sassone, “Riskio: A Serious Game for Cyber Security Awareness and Education,” *Computers & Security*, vol. 95, p. 101827, Aug. 2020, doi: 10.1016/j.cose.2020.101827.
- [102] C. Maze, A. Haye, J. Sarre, M. Galaup, P. Lagarrigue, and C. P. Lelardeux, “A Board Game to Fight Against Misinformation and Fake News,” in *Games and Learning Alliance*, Cham, 2020, pp. 326–334. doi: 10.1007/978-3-030-63464-3_31.
- [103] R. Maertens, J. Roozenbeek, M. Basol, and S. van der Linden, “Long-term effectiveness of inoculation against misinformation: Three longitudinal experiments.,” *Journal of Experimental Psychology: Applied*, 2020.
- [104] Fred and Fred, “Cambridge game ‘pre-bunks’ coronavirus conspiracies,” *University of Cambridge*, Oct. 11, 2020. <https://www.cam.ac.uk/stories/goviral> (accessed Mar. 25, 2021).
- [105] G. A. Bowen, “Document Analysis as a Qualitative Research Method,” *Qualitative Research Journal*, vol. 9, no. 2, pp. 27–40, Jan. 2009, doi: 10.3316/QRJ0902027.
- [106] A. R. Hevner, S. T. March, J. Park, and S. Ram, “Design Science in Information Systems Research,” *MIS Quarterly*, vol. 28, no. 1, pp. 75–105, 2004, doi: 10.2307/25148625.
- [107] T. W. Edgar and D. O. Manz, *Research methods for cyber security*. Syngress, 2017.
- [108] S. Bell, “Design thinking,” *Temple University Libraries*, 2008.

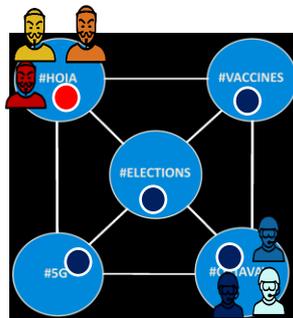
Appendix 1

Hashtag Struggle Rules Summary

Hashtag Struggle is a game for six players and a facilitator. The facilitator assigns each player an actor, their six units, and objectives according to the following table. This designation should be done in a way guaranteeing the objectives are revealed only to the respective player.

Actor, team, and objective	Units	Order of play
BIA, Blue team. Objectives: Protect the following communities: #VACCINES, #ELECTIONS, and #5G.	 ¹	1
FNF, Red team. Objectives: Have 2 units in the following communities: #5G, #VACCINES, and #HOIA	 ²	2
FB, Blue team. Objectives: Protect the following communities: #OSTAVAY and #VACCINES		3
HPP, Red team. Objectives: Distort the following communities: #HOIA and #OSTAVAY		4
FWB, Blue team. Objectives: Find HPP's communities. Have at least 2 units in each.		5
AIR, Red team. Objectives: Distort the following communities: #VACCINES, #ELECTIONS, and #HOIA.		6

The Blue team and all their units start in the #OSTAVAY community. Similarly, the Red team begins in #HOIA with all their units. In addition, all communities except #HOIA should have their token status set to protected. The following images shows the board of the game at the start.



¹ "soldier" by icon 54, used under CC BY / Recoloured from original.

² "anonymous" by icon 54, used under CC BY / Recoloured from original.

Hashtag Sequence of Play

Following the player order stated in order of play column, each player executes a single action during their turn. A round ends when all players have executed an action. The game ends after 30 rounds.

Hashtag Struggle Player Actions

Bridge. Create a new connection between two communities. The player must have units present in the community from where the connection originates. The player choosing this action must roll two six-sided dice. If the result is greater than or equals to 7, the connection is created.

Narrow. Remove an existing connection between two communities. The player must have units present in the community where the connection will be removed. It cannot be used to remove the last remaining connection. The player choosing this action must roll two six-sided dice. If the result is greater than or equals to 7, the connection is removed.

Exchange. Move or restore any number of player-owned units. Units move between two connected communities. When restoring units, they must be placed in communities with friendly units.

Engage. Engage allows the removal of opposing units in a specific community where the attacking player has units. When engaging, the attacker states the community and the player acting as the defender. Both the players roll a six-sided dice per unit where the engagement is occurring. The dices are paired by taking the highest value from each side until no more dice pairs can be formed. Within each dice pair the values are compared, with higher values defeating lower ones. For each defeat in a pair, the losing side must remove a unit from the community. Any unpaired dice are ignored.

Hashtag Struggle Key Concepts

Unit. The pieces controlled by each player. Each player has a total of six units available.

Community. The terrain of Hashtag Struggle. There are five communities in total.

Connection. The bridge between two communities. It determines valid movement options for the players' units. Players can add or remove connections.

Protected. A community is protected when the total number of blue units is greater than the total number of red units.

Distorted. A community is distorted when the total number of red units is greater than or equal to the total number of blue units.

Hashtag Struggle Videos

The following playlist shows a gameplay video.

https://youtube.com/playlist?list=PLajd_8PdQTzHQUpYbNGPxVpsN4mxDkW9V

Note: The participants provided written consent to allow the sharing of this video.

Appendix 2

Pre-session questionnaire

Information manipulation on social media

This anonymous questionnaire is to support a master thesis on the use of serious games for cybersecurity training on the topic of information manipulation on social media.

1. What gender do you identify as?

Mark only one oval.

- Female
- Male
- Prefer not to say
- Other: _____

2. What is your age?

Mark only one oval.

- Under 12 years old
- 12 - 17 years old
- 18 - 24 years old
- 25 - 34 years old
- 35 - 44 years old
- 45 - 54 years old
- 55 - 64 years old
- 65 - 74 years old
- 75 years or older

3. Concerning your education and training, how have they been delivered? Select all those that apply:

Check all that apply.

- Instructor/lecturer-led
- eLearning/e-course/web-based
- N
- Group discussions and activities
- Coaching or mentoring

Other: _____

4. Which of the following concepts are you familiar with? Select all those that apply:

Check all that apply.

- Disinformation
- Fake News
- Misinformation
- Malinformation
- None of them

5. What word would you use to refer to intentionally misleading, false, or deceptive information?

Mark only one oval.

- Malinformation
- Fake News
- Misinformation
- Disinformation
- None of them
- Other: _____

6. What word would you use to refer to unintentionally misleading, false, or deceptive information?

Mark only one oval.

- Malinformation
 Fake News
 Misinformation
 Disinformation
 None of them
 Other: _____

7. Which of the following cybersecurity concepts are you familiar with? Select all those that apply:

Check all that apply.

- Confidentiality
 Integrity
 Availability
 Risk
 Adversarial Thinking
 System Thinking
 Not aware of any of the previous concepts

Other: _____

8. In 1 to 3 sentences, what do you think 'adversarial thinking' is about?

9. How confident are you this is the right answer?

Mark only one oval.

	1	2	3	4	5	
Not confident at all	<input type="radio"/>	Very confident				

This content is neither created nor endorsed by Google.

Google Forms

Appendix 3

Pre-session questionnaire results

What gender do you identify as?	What is your age?	Concerning your education and training, how have they been delivered? Select all those that apply:	Which of the following concepts are you familiar with? Select all those that apply:	What word would you use to refer to intentionally misleading, false, or deceptive information?	What word would you use to refer to unintentionally misleading, false, or deceptive information?	Which of the following cybersecurity concepts are you familiar with? Select all those that apply:	In 1 to 3 sentences, what do you think 'adversarial thinking' is about?	How confident are you this is the right answer?
Male	25 - 34 years old	Instructor/lecturer-led;eLearning/e-course/web-based;Group discussions and activities;Coaching or mentoring	Disinformation;Fake News;Misinformation;Malinformation	Malinformation	Misinformation	Confidentiality;Integrity;Availability;Risk;Adversarial Thinking;System Thinking	Putting yourself in the shoes of the person who's going to pen test your system for weaknesses/modelling the sort of attacks your system can expect to see	4
Male	18 - 24 years old	Instructor/lecturer-led;eLearning/e-course/web-based;Group discussions and activities	Disinformation;Fake News;Misinformation	Disinformation	Misinformation	Not aware of any of the previous concepts	I assume it is linked with assuming someone's knowledge in some field. Not sure though, second time in my life that I hear this expression.	2
Female	18 - 24 years old	Group discussions and activities	Fake News;Misinformation	Misinformation	Malinformation	Confidentiality;Integrity;Risk	Hackers could maybe adapt the specific way of thinking of another, to copy ones strategy.	2
Male	25 - 34 years old	Varies.	Disinformation;Fake News;Misinformation;Malinformation	Disinformation	Misinformation	Confidentiality;Integrity;Availability;Risk;Adversarial Thinking	Knowing what to expect from cyber attackers & adversarial party	5
Male	18 - 24 years old	Instructor/lecturer-led;eLearning/e-course/web-based;Game-based learning;Group discussions and activities;Coaching or mentoring	Fake News;Misinformation	Malinformation	Disinformation	Adversarial Thinking	Thinking and thereby seeing things from another angle. This way helps to	5

							understand the motives of another person. For instance, a detective should try to see things the perspective of a person who committed a crime, this way it helps to identify the person, and what his motives could be.	
Female	18 - 24 years old	Instructor/lecturer-led	Disinformation;Fake News;Misinformation	Disinformation	Misinformation	Confidentiality;Availability;Risk	The meaning of the concept is the capability to think for a step further like a hacker	2
Male	18 - 24 years old	Group discussions and activities	Disinformation;Fake News	Disinformation	Misinformation	Confidentiality;Integrity;Risk	Some action against somebody else, thought from the perspective in how this other would act	3
Male	25 - 34 years old	eLearning/e-course/web-based;Group discussions and activities	Fake News;Misinformation;Malinformation	Disinformation	Misinformation	Confidentiality;Integrity;Availability;Risk		4
Male	25 - 34 years old	Instructor/lecturer-led;eLearning/e-course/web-based;Group discussions and activities	Fake News;Misinformation	Fake News	None of them	Confidentiality;Integrity;Availability	Thinking like the hacker	5
Female	25 - 34 years old	Instructor/lecturer-led;eLearning/e-course/web-based;Group discussions and activities;Coaching or mentoring	Disinformation;Fake News;Misinformation	Fake News	Misinformation	Confidentiality;Integrity;Risk	To think like the person we are performing an attack or we are securing from.	3
Male	25 - 34 years old	eLearning/e-course/web-based	Disinformation;Fake News;Misinformation	Disinformation	Misinformation	Confidentiality;Integrity;Availability;Risk	Trying to think in a similar way like a cyber criminal.	3
Male	25 - 34 years old	Instructor/lecturer-led;eLearning/e-course/web-based;Group discussions and activities;Coaching or mentoring	Disinformation;Fake News;Misinformation	Disinformation	Misinformation	Confidentiality;Integrity;Availability;Risk;System Thinking	Adversarial thinking is a pattern of thinking that contradicts and challenges trend and narrative.	2

Appendix 4

Post-session questionnaire

Information manipulation in social media

This anonymous questionnaire is to gather feedback on the piloting session.

1. What gender do you identify as?

Mark only one oval.

- Female
 Male
 Prefer not to say
 Other: _____

2. What is your age?

Mark only one oval.

- Under 12 years old
 12 - 17 years old
 18 - 24 years old
 25 - 34 years old
 35 - 44 years old
 45 - 54 years old
 55 - 64 years old
 65 - 74 years old
 75 years or older

3. I thought the game was fun:

Mark only one oval.

	1	2	3	4	5	
Strongly disagree	<input type="radio"/>	Strongly agree				

4. When playing the game, I felt I was learning:

Mark only one oval.

1 2 3 4 5

Strongly disagree Strongly agree

5. I found the debriefings to be useful:

Mark only one oval.

1 2 3 4 5

Strongly disagree Strongly agree

6. How can the game be improved? Select all those that apply

Check all that apply.

- Shorter playing time
- Longer playing time
- More engaging / entertaining / fun
- Better / more educational material
- Clearer / simpler / easier instructions and rules
- More appealing/higher quality game components

Other: _____

7. In your own words, define adversarial thinking:

8. How confident are you this is the correct definition?

Mark only one oval.

1 2 3 4 5

Not confident at all is the correct definition Very confident it is the correct definition

9. What is your main reflection from the whole experience?

10. In a scale from 1 to 10, how would you rate the game?

Mark only one oval.

1	2	3	4	5	6	7	8	9	10
<input type="radio"/>									

11. In a scale from 1 to 10, how would you rate the piloting session (game and debriefings)?

Mark only one oval.

1	2	3	4	5	6	7	8	9	10
<input type="radio"/>									

12. Would you like more serious games to be incorporated in your education and training?

Mark only one oval.

	1	2	3	4	5	
Not at all	<input type="radio"/>	Very much				

This content is neither created nor endorsed by Google.

Google Forms

Appendix 5

Post-session questionnaire results

What gender do you identify as?	What is your age?	I thought the game was fun:	When playing the game, I felt I was learning:	I found the debriefings to be useful:	How can the game be improved? Select all those that apply	In your own words, define adversarial thinking:	How confident are you this is the correct definition?	What is your main reflection from the whole experience?	In a scale from 1 to 10, how would you rate the game?	In a scale from 1 to 10, how would you rate the piloting session (game and debriefings)?	Would you like more serious games to be incorporated in your education and training?
Male	25 - 34 years old	5	5	5	Longer playing time;Clearer / simpler / easier instructions and rules	Thinking through the lens of someone whose interest might be to cause you harm.	4	Fun, Interesting and informative	7	8	4
Male	25 - 34 years old	5	4	5	Let Mauricio make a game	Game theory.	5	Mauricio has the patience of a saint. The game is also legitimately fun, and would be more fun were it not for the constraints of academia	8	10	1
Male	25 - 34 years old	4	4	5	More engaging / entertaining / fun;More actions would be preferred. Also, some kind of a concrete ending point for the game would make it more engaging.	You try to think like a cyber criminal and what a cyber criminal would do.	4	A great effort has been taken to plan the game. At first it was somewhat confusing what is the aim of the game, but later on it became more clearer. It was understandable what the game is trying to teach, but it would be better if there was more actions than move troops, cut or bridge lines and engage in a fight. Also,	8	9	5

								would be good if there was a definite ending point for the game.			
Female	25 - 34 years old	5	3	5	Clearer / simpler / easier instructions and rules; Clear overall goal, example does the game end when 3 spaces are covered or 4?	To think like a hacker, to understand what and how vulnerabilities can be attacked.	4	That in terms of creating fake media, noise or volume makes a difference. The more people talking about it, the more it gets spread. Also the game taught us about how communities are interconnected with each other and hence that can make an impact in the communication outcome.	8	7	5
Male	18 - 24 years old	4	3	3	Clearer / simpler / easier instructions and rules	To think how the other will react in future decisions and act accordingly to have influence over the other person or community.	3	It was pretty fun, not sure the final purpose of the Master Thesis, but the game was fun.	7	7	4
Male	18 - 24 years old	5	2	3	Longer playing time; Clearer / simpler / easier instructions and rules; As have been mentioned in the call, it would be wise to think about the roles of each player's pawns, what advantages/disadvantages they might have when being on a specific circle. Additionally, it would be also better to think about what advantages might circles give to a player/team when it is under someone's control.	Basically trying to think about how other person might act (behave). I would not say that it's a "cybersecurity principle", I believe, it is a universal principle used in many different fields - criminology, psychology, as well as playing other games like chess, when you have to think in advance what might be the next step of the opponent.	5	The game is very fun, would be wise to finalize it and commercialize it.)	9	8	5
Male	25 - 34 years old	5	4	4	Clearer / simpler / easier instructions and rules	What to expect from your adversary & from an attack.	5	Part of the appeal was not to know what to expect. The strategical elements were present, point comes clear and overall the experience and communication was fun. The "game" elements might need some fine-tuning.	9	10	5
Female	18 - 24 years old	5	4	5	Longer playing time; Better / more educational material	The capability to comprehend atypical perspectives and tactical thinking of hackers	4	The game reminded me of the reality when the society needs to work in the form of a team, for instance, to work together for achieving the UN Sustainable Development Goals (in the game it was represented by the teams of	9	9	5

								red and blue, who were required to disinform and protect the fields respectively), however, each of the states involved in the common system has different objectives to reach for its own advantage (in the game each player had different objectives to achieve).			
Male	25 - 34 years old	5	4	5	More engaging / entertaining / fun;Clearer / simpler / easier instructions and rules	Thinking like the attacker/hacker who might want to mess with your system	5	It was educational and interesting!	8	10	5
Male	18 - 24 years old	5	5	5	Longer playing time;More engaging / entertaining / fun	"Thinking like a hacker". Basically it means thinking ahead and trying to predict the abilities of the other person.	3	It was very interesting, fun and educational experience.	10	10	4
Male	25 - 34 years old	1	1	2	Better / more educational material;Clearer / simpler / easier instructions and rules		2	little bit confusing like in real life	3	4	3
Female	18 - 24 years old	5	5	5	Longer playing time;Better / more educational material;More appealing/higher quality game components	Adversal thinking is about being ahead one step of others, continuously trying to anticipate the strategic or even big entities.	4	Overall the piloting session experience was a top notch, something that would definitely be useful to combine into university studies. From the perspective of a lawyer, this kind of thinking is truly important and also overall in a fast developing world we need to adapt to new things, be ready to adapt new patterns in a multicultural environment. The educational lesson of the game also reached us, the players and personally I learned a lot new things. The element of fortune caused by the dices was also creative, as nothing in life can be straight forward calculated. All the best to you Mauricio and good luck with the thesis! :)	9	10	5