systems

MDPI

Article

Decision-Makers' Understanding of Cyber-Security's Systemic and Dynamic Complexity: Insights from a Board Game for Bank Managers

Sander Zeijlemaker [1,2,*], Etiënne A. J. A. Rouwette [3], Giovanni Cunico [4], Stefano Armenia [5] and Michael von Kutzschenbach [6]

1 Cybersecurity at MIT Sloan, Sloan School of Management, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
2 Disem Institute, 1827 LR Alkmaar, The Netherlands
3 Institute for Management Research, Radboud University, 6525 AJ Nijmegen, The Netherlands; etienne.rouwette@ru.nl
4 Business School, University of New South Wales (UNSW), Sydney, NSW 2052, Australia; g.cunico@unsw.edu.au
5 Department of Research, Link University, Via del Casale di San Pio V, 00165 Rome, Italy; s.armenia@unilink.it
6 Institute of Management, University of Applied Sciences and Arts Northwestern Switzerland, 4002 Basel, Switzerland; michael.vonkutzschenbach@fhnw.ch
* Correspondence: s.zeijlemaker@disem-institute.com

Abstract: Cyber-security incidents show how difficult it is to make optimal strategic decisions in such a complex environment. Given that it is hard for researchers to observe organisations' decision-making processes driving cyber-security strategy, we developed a board game that mimics this real-life environment and shows the challenges of decision-making. We observed cyber-security experts participating in the game. The results showed that decision-makers who performed poorly tended to employ heuristics, leading to fallacious decision approaches (overreaction strategies in place of proactive ones), and were not always aware of their poor performances. We advocate the need for decision support tools that capture this complex dynamic nature.

Keywords: cyber-security dynamics; cyber-security management; cyber-security board game; banking cyber-security; cyber-security training; system dynamics gamification

Citation: Zeijlemaker, S.; Rouwette, E.A.J.A.; Cunico, G.; Armenia, S.; von Kutzschenbach, M. Decision-Makers' Understanding of Cyber-Security's Systemic and Dynamic Complexity: Insights from a Board Game for Bank Managers. *Systems* **2022**, *10*, 49. https://doi.org/10.3390/systems10020049

Academic Editor: William T. Scherer

Received: 7 March 2022
Accepted: 6 April 2022
Published: 14 April 2022

1. Introduction

Limited access for researchers to boardrooms and the unpredictability of cyber-attacks makes it hard to consistently observe decision-making concerning cyber-security over extended periods in organisations and companies. Nevertheless, managerial and financial cyber-security decisions play a crucial role in modern companies, as they can have significant organisational and economic impacts [1,2]. This is evident in several recent real-life cyber-security incidents, e.g., WannaCry [3], Carbanak [4], and Diginotar-hack [5], where inadequate resource allocation led to security breaches. Such incidents may then affect companies' competition [6] and market value [7]. Decisions on cyber-security are not only relevant to the security system but also to the overall behaviour of the organisation. However, making decisions in such contexts is difficult, as they confront decision-makers with a complex and dynamic system [8]. Specifically, the underlying systemic conditions constantly change over time and are driven by intricate and delayed interdependencies among the various areas of the company, due to the presence of feedback loops [8–10]. The literature reports various explanations for poor managerial decisions in cyber-security. These range from the misalignment of interests between different internal [11] and external [12] stakeholders to economic arguments (striking for a balance between over- and underinvestments [13]) and willingness to pay [14]; from social and

behavioural ones (e.g., 'cover-my-ass-security' strategy, when a lot of resources are spent on a security issue by managing defenders to show to have done their best and protect themselves from future criticism [15]) to just psychological ones, such as fear, uncertainty and doubt [16]. Overall, these reasons can be traced back to the challenges of understanding and managing complex systems [17–20].

Given the limitations to the empirical observation that we mentioned above, other researchers, such as Moore et al. [21], used nonrecurring (semi)structured interviews to study the cyber-security managerial decision-making process. According to this research, there is a strong focus on bridging the difference between the current and the desired security posture as well as on the process of decision-making. However, the interview data do not fully capture how the complex and dynamic nature of the underlying security system is considered in the strategic decision process, possibly compromising the effective understanding of the phenomena at stake. Serious games are used in system dynamics to investigate the behaviour of complex dynamic systems.

At the same time, serious games are increasingly being used in the field of cyber-security, both as training tools and as instruments to encourage security behavioural change [22–24]. At the core of many of such "gaming simulations", there is the interaction between attacker and defender [24,25], mainly focusing on the IT technical side, namely on infrastructure (e.g., network security), defensive procedures and staff awareness (e.g., phishing countermeasures). Surprisingly, cyber-security management decisions and related financial consequences in a business context are often not fully considered (or even completely disregarded) in such games [22].

Therefore, we developed a cyber-security board game (called Cyber-Security Game: 'Red versus Blue') that mimics a real-life strategic cyber-security decision-making environment. The objective was to observe how cyber-security managerial decisions are taken within an organisation and improve cyber-security managers' and practitioners' awareness and understanding of the financial consequences of their actions. By analysing players' performance during several gaming sessions and by delving into players' in-game observations and ex post evaluations, we tried to explore and improve the comprehension of decision-making processes in these complex dynamic cyber-security systems.

This research answers the call to improve the managerial decision process [26,27] in both a practical and theoretical sense. First, it describes and reports on a newly developed tool that captures the main elements of the decision-making process in a complex dynamic cyber-security environment. Second, it analyses the results of several gaming sessions and reflects on them so as to gain general insights into the decision process. We believe that these insights will ultimately contribute to and support cyber-security managers in their decisional tasks by highlighting potential pitfalls in the decision processes.

The paper is structured as follows: Section 2 briefly describes the theoretical background on which this research relies; Section 3 extensively reports on the underlying game logic, its relation to reality and the translation into a board game; Section 4 presents and analyses the data collected during game sessions; Section 5 discusses these results and presents some of the main insights; Section 6 reflects on the limitations and further research directions; and Section 7 summarises and generalises our conclusions.

2. Background

Research shows that people experience difficulties when making decisions in complex dynamic environments [17,28], such as in the management of cyber-security, and tend to use simple mental rules called "heuristics" [17,18,20]. Heuristics can be defined as "adaptive tools that ignore information to make fast and frugal decisions that are accurate and robust under conditions of uncertainty" [29]. We can also say that heuristics are practical approaches to problem-solving in complex situations, far from being ideal or rational ones [30–32]. In fact, most of the time, they do not lead to optimal decisions [17,18,20,33–35]. However, heuristics are usually adequate for achieving immediate, short-term objectives or approximations [30–32], which may even be sufficient in cases in which optimal solutions

are hard to find. Moreover, as heuristics aim at expediting the development of a decision, they offer mental shortcuts that decrease decision-maker cognitive burden [30–32]. Heuristics are generally based on previous individual experiences and may employ, for example, trial and error mechanisms or intuitive decisions (i.e., estimations and conjectures based on the prior knowledge and experience of a decision-maker that are derived from intuitive control [36]). For these reasons, managers often tend to base their decisions on heuristics when dealing with complex and uncertain contexts [30]. Although decisions based on heuristics tend to work in the short term, they may lead to significant undesired side-effects. Examples of possible consequences include an unintended and unconscious overestimation or underestimation of the actual situation [30], an inadequate assessment of the decisions' effects on peers [17] or a decreasing attention to repeated similar informational inputs over time [37]. Cyber-security managers, just as all decision-makers, also tend to use heuristics [38–40]. For instance, research shows that cyber-security decision-makers' background and experience deeply influence their judgement and risk perception [38,40] by eventually pushing them to create misleading correlations between the past and the present to develop defensive strategies [39], although the two may be significantly different. This can lead to inadequate defense arrangements and plans (e.g., under- or overestimation of threats) or poor resource allocation in terms of quantity and timing, thereby increasing the chance of future successful attacks, ultimately generating two types of costs for the organisation: (a) necessary investments in new security infrastructures and (b) mitigation and restoration costs [2]. In this light, Jalali et al. [41] highlight that cyber-security professionals do not perform better than inexperienced decision-makers in the same cyber-security game setting.

The practical and theoretical considerations regarding the difficulties of observing cyber-security decision-making processes and the increasing recognition of serious games as relevant experimental tools led us to a research question that, in the first instance, aims at assessing the usefulness of the developed game to observe players' decision-making processes. Secondly, we deliberated as to what decision-making practices could be observed from the players' behaviours in the game's sessions, contributing to the recent debate on the employment of heuristics in cyber-security [38–40].

3. Method

In order to develop Cyber-Security Game: 'Red versus Blue', we first built a system dynamics (SD) simulation model to capture the complex dynamics in which cyber-security managers make decisions. Then, the model was transferred into the rules and mechanics of a board game that replicated the context of an internet bank. It is worth noting that a board game can be basically seen as a consequence-free environment that allows players to test the outcomes of their decisions and also allows us to observe the process and nature of players' decision-making during each game. The fact that the game was based on an SD model ensured that the incorporation of the cyber-security resource allocation dynamics and subsequent system consequences had a degree of robustness and was in line with real-life decision-making.

This section will introduce and explain the SD method, how it was used to build a model describing the system of interest and, finally, how the model was used as a basis for an interactive learning environment, the Cyber-Security Game: 'Red versus Blue' board game.

3.1. System Dynamics

In this paper, we explore potential synergies to develop a new SD-based cyber-security board game: 'Cyber-Security Game: 'Red versus Blue', building on the lessons gained from research on serious games and cyber-security.

SD is a modelling approach used to describe, simulate and analyse dynamically complex issues in terms of their processes, interdependencies, information, organisational boundaries and strategies. Sterman [28] defines SD as a particularly useful modelling and simulation method when the goal is to analyse sociotechnical nonlinear complex systems. SD

aims at capturing the underlying structure driving system behaviour [42,43]. SD modelling also results in learning about dynamic complexity—caused by feedback/interdependencies, accumulation and time delays—and in understanding the sources of policy resistance, which helps to overcome them and ultimately promote the design of more effective policies [28,44,45].

There is a long and robust tradition of SD-based serious games, both single- and multiplayer, and physical and virtual, [46–48] that serve the purpose of improving the understanding of systems (either by players through exploration and training or by researchers through observation of players' decisions). 'Stratagem2' [43], the Beer Game [18] and the Sustain game [49] are examples of well-known successful SD-based board games. Research indicates that players' interaction with serious games, particularly with board games, has several positive impacts, such as a deeper understanding of content, behavioural changes, knowledge retention and soft-skills development [23,50,51]. On top of this, SD-based games also aim to improve players' decision-making skills in dynamic and complex contexts [44,47,51]. These skills are necessary for managing both complex dynamic environments [28] and for tightly coupled ones, such as critical infrastructures (which constitute the context in which most cyber-security decisions occur [41]).

SD has, in fact, recently started to be used to study cyber-security decision-making processes. For instance, recent studies have addressed decision-making related to insider threats [52], cooperation and learning during incident responder training programs [53], the dynamic and systematic assessment of cyber-security risks [54], and investment decisions [41]. Our research is the first attempt to focus on strategic cyber-security management.

3.2. System Analysis

As mentioned above, cyber-security decisions are usually taken in a complex organisational environment and generally involve a broad range of managers [8–10,55]—such as IT, business operations, security, risk and finance—and affect different stakeholders [1], such as threat actors, customers, supervisors and legislators. This introduces the potential for highly counterintuitive and delayed behaviours, mostly arising from multiple agents' interactions over time and the fact that agents have different intrinsic goals [41,42,56,57]. From a high-level perspective, cyber-security decision-makers need to balance their security strategies between response (aimed at managing unpredicted threats. Threats are 'any circumstances or events with the potential to cause harm to any information resource by exploiting vulnerabilities in the system' ([58], p. 108). A vulnerability is 'a weakness in the design, implementation, operation or internal controls in a process, that can be exploited to validate system security' ([58], p. 287)) and prevention (aimed at managing predicted threats) [59]. They are forced to find the optimal allocation of resources (funds, staff, etc.) between the two extreme orientations of these strategies: on the one hand to ensure an ideal level of security from cyber-incidents (i.e., an observed attack) and, on the other hand, to minimise the costs (i.e., security investment) for the organisation. However, the environment in which decision-makers have to make those decisions is dynamic, and marked by two recurring systemic structures [8,9].

First, an important characteristic driving the behaviour of the system is the interaction between attacker and defender. This can be seen as a (constantly changing) struggle between the two parties to find/solve weaknesses in the defense perimeter. The attacker tries to exploit weaknesses for a successful attack, and the defender attempts to block such attacks [60]. Both defender and attacker rely on a learning curve based on their past performances in order to anticipate the other's actions and adjust their strategy accordingly [34]. After a successful attack, defenders tend to upgrade their defense measures when they are uncertain about the effectiveness of their future defensive actions [60], or the attack's impact exceeds the defender's level of impact acceptability [61]. Notably, this whole structure is represented as a feedback loop in the system, i.e., a variable's output circles back through a set of cause–effect relationships, becoming an input to the variable. In this case, defender capability decisions depend on the outcomes of an attack, which are

in turn affected by the defender's capabilities. An equivalent and opposing process is also in place for the attacker, as attackers learn from the outcomes of their attack and use the information to upgrade their capabilities. Consequently, attacker and defender are in a constant, tightly coupled and escalating 'arms race' of additional capabilities and resources.

Second, there is a response from the organisation aimed at achieving an ideal level of security from cyber-incidents. This mechanism captures how the defender tries to anticipate upcoming attacks, as well as how they try to learn from cyber-incidents that have already occurred. In the field of cyber-security, organisational learning can be proactively facilitated by, among other things, threat intelligence (Threat intelligence reports detailed information about potential cyber-threats with insights on network structure, operations and activities [62]. Such an alerting system discloses vulnerabilities at an early stage and enables the defender to anticipate threats early on [63].), runbooks (A runbook is the written collection of procedures and operations that a system administrator, operator or security specialist executes when an emergency occurs.), and exercises (Exercises are simulated attacks that mimic cyber-attacks with the purpose of testing and improving the defences in place.) [8,9], with the result that security capabilities are better aligned to upcoming cyber-attacks. In this way, the organisation prepares for improving current and future defences to enhance business stability and performance following potential future cyber-incidents and disruptions. Through learning, resilience emerges, and the organisation becomes strengthened and more resourceful [64]. Resilience is determined by the speed at which a breached organisation is brought back to its normal state after a security incidents [65]. What characterises the resilience of organisations is their superior organisational learning capabilities [64,66], a quality particularly needed for contemporary organisations to adapt their capabilities to handle future security threats.

In this context, it is argued that each action taken by decision-makers depends on financial considerations and has economic consequences, which are bounded by and to the financial capabilities of the defender [13]. The complex challenge for cyber-security decision-makers is to find an optimal solution that serves both low costs and a high cyber-security level (i.e., decreased risks for the organisation). It is argued that investing in the right security strategies and anticipating the attacker's actions through learning are necessary conditions.

In response to such complex tasks and decisions, cyber-security managers may be tempted to use heuristics. However, the reliance on simplistic problem-solving methods when dealing with complex problems may be risky for the organisation, particularly regarding cyber-security defences. For example, let us suppose that a company's cyber-security managers poorly evaluate the state of the system: they estimate a low probability and impact for specific threats. To minimise costs, they do not update the organisation's defence capabilities (i.e., executing the correct prevention strategy), while the adversaries continue to evolve and the cyber risks increase imperceptibly. This is a likely scenario for managers who base their decisions on previous short-term experiences (e.g., no recent attacks) or trial and error (two typical heuristic approaches [30–32]). As a consequence, more attacks towards the organisation will be successful and more resources will be needed to later mitigate the damages of the incidents (i.e., response strategy), potentially leading the organisation into the capability trap [35]. A "capability trap" occurs when decision-makers require increasing levels of effort to sustain performance "instead of making up for the improvement activity they skipped earlier" [35], p. 282. In the case of continuous successful attacks, management is faced with the need to reduce the resulting attack's impact (response strategy). As such, they are caught "in a downward spiral of eroding process capability, increasing work hours and less and less time for improvement" [35], p. 282. Specifically, cyber-security managers will not have the necessary resources to invest in prevention strategies that would benefit the organisation in the short and long term, since such resources will have to be directed towards the current response to the incidents.

3.3. Model Development and Validation

The previously described dynamics were captured using the SD modelling and simulation approach, and the resulting model is depicted in Figure 1, below. The model took a defender perspective, as this was the point of view that was of interest to us, and was rooted in the literature and systems analysis described in the previous sections. After a first draft, the model was refined with the help of cyber-security specialists and managers, in particular by involving them in the model construction process by using a well-known participatory modelling approach known as Group Model Building (GMB) [19]. This method consists in using facilitation techniques to extrapolate, and capture into SD terms, the knowledge of a group of experts. Besides improving the participants' systems understanding and achieving a consensus on policies [67], GMB also helps to improve the resulting model's reliability and quality [19], as participants can immediately have a first-hand experience with the 'real' system [68]. In this study, we held three participatory GMB modelling sessions involving two security consultants, a business consultant, a training and awareness specialist and a team manager. All of these participants were from the cyber-security department of a large European corporation. It is worth noting that the wide range of expertise among the GMB participants generally provided further reliability to the model's robustness [19].

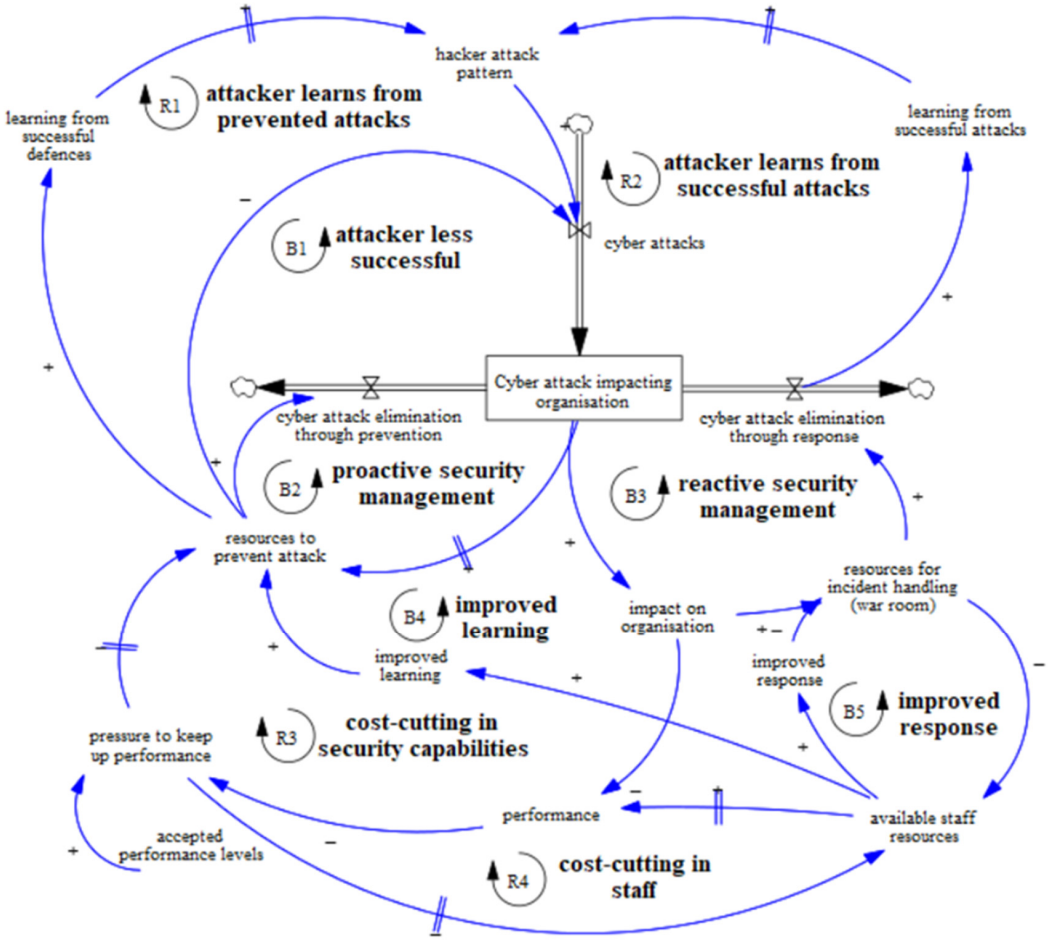

Figure 1. Underlying aggregated systemic structure of security investment decision-making.

Figure 1 shows the result of the modelling efforts by means of a causal loop diagram. First, we explain the syntax used in this figure and thereafter the structure of the model this figure represents. A causal loop diagram derives from drawing, in this case through the GMB participatory approach previously described, the various "causal" relationships that exist among those aspects that the experts considered as most relevant. A causal relationship between two events exists if the first event can be determined as the cause of the second

one, whereas the latter can be considered as the effect triggered by the former. A closed loop composed of various causal relationships is called a "feedback loop". Such feedback "mechanisms" are visible in Figure 1 and are labelled by the letters "R" and "B", as well as numbered from one to five. Feedback between variables in a system can be reinforcing ("positive") or balancing ("negative"). Reinforcing feedback mechanisms (labelled with R) are characterized by exponential growth and hence may act as an engine driving the general growth of certain variables in the model. Balancing feedback mechanisms (labelled with B) instead evoke goal-seeking behaviour. Reinforcing means that two variables in the system strengthen each other. Balancing means that two variables in the system have an oppositive effect on each other. Figure 1 also shows the presence of time delays (evidenced by the symbol "||"). A time delay is the time-dependent effect of changing the feedback in a system. Time delays mean that the feedback effect of one variable on another variable is not instantaneous and will take some time to become impactful.

To summarize, Figure 1 portrays the aggregated view of our model and shows that the attacker learns from both unsuccessful ($R1$) and successful attacks ($R2$) and can improve its future attack actions. The defender can proactively prevent this ($B2$), which depends on the security resources and capabilities in place, or mitigate the effects of the attack through responsive actions of security incident handling ($B3$). These responsive actions will impact the organisation and require additional staff allocation to enable security incident handling. These feedback mechanisms are relevant to the attacker–defender interaction. Staff can also work on improvements that foster learning ($B4$) and response quality ($B5$), which allow the defender to react better to the attacker's actions and, accordingly, to react more effectively to incidents. If financial performance is impacted and falls below accepted levels, cost cutting is needed in the staff ($R3$) and/or security capabilities ($R4$). These feedback mechanism are relevant to the (dis)appearance of the capability trap. In this model, all reinforcing (R) feedback mechanisms contribute to the strength of the attacker and all balancing feedback mechanisms (B) contribute to the strength of the defender.

Following a first qualitative description, the presented causal-loop diagram was populated with quantitative data defining each variable and each relationship and was then simulated. The simulation output replicated the observed patterns of the central variables in the issue [68], thus allowing the model to pass the "behaviour reproduction" validation tests [28,69,70]. The model was also validated through structural evaluation tests: the correctness of the equations used was compared against the empirical knowledge and scientific literature about the real system [28,69,70]. In three group modelling sessions, the model's structure was created using knowledge from experts in the field [55]. Furthermore, a different group with similar expertise was involved in two participatory sessions to assess the model [55]. Part of this process was a model walkthrough based on cyber-security incidents that have occurred in real life. Their approval provided further validation and provided us with sufficient confidence in its ability to capture strategic cyber-security management. More details on the model construction, structure, simulations and validation can be found in [55].

3.4. Game Design

This section outlines how the model principles that we described above were translated into a cyber-security board game. Specifically, we contextualized the model structure to an online bank by creating an ad hoc background narrative, also adapting the variable names and objects to the game context and establishing performance and winning criteria fully in line with a typical bank organisation. The reason for this choice was twofold: theoretical and practical. The theoretical reason was that evidence of the importance of effective and safe IT components in banks' corporate strategy has been gathered for quite some time now [71,72], which is also recognized in specific laws and regulations [73,74]. The practical reason was that all the participants in the various game sessions were working in the cyber-security sector of a bank, thus making it more straightforward to incorporate into the game situations that they were familiar with. However, we strongly believe that the

underlying model structure can also be adapted to other types of organisations dealing with strategic cyber-security management (which will be one of the goals for future research applied in this context).

The rest of this section is devoted to explaining the game's setting, the game's round sequence, the game board design, the winning criteria and the game validation.

3.4.1. Setting

As previously mentioned, the game depicts the cyber defence of an internet bank. The game includes six roles (five players, plus a facilitator).

- Four Bank Defenders (blue team):
    1. The Security Engineer (SecEng) is responsible for selecting, investing, maintaining and decommissioning security capabilities to protect the bank from specific attacks;
    2. The Development and Operation (DevOps) manager is responsible for maintaining the business operations supporting customer transactions as well as responding to incidents;
    3. The Cyber Emergency Response Team (CERT) manager is responsible for reacting to cyber-security incidents by proactively analysing threat intelligence and improving responsive policies, or reactively responding to incidents;
    4. The Chief Executive Officer (CEO) is responsible for budget allocation, gaining and maintaining customers, reporting on cyber-incidents and the financial performance of the internet bank.
- One Attacker (red team's only player): this player analyses the blue team's behaviour and selects a set of attack playing cards. These cards represent the attacks on the internet bank.
- The Facilitator: this role oversees data gathering, supports the game process and explains the rules to the players. The facilitators received ten hours of training.

Figure 2 shows the game roles along with their actions and is closely related with the game data collected per round. Table 1 shows the collected game data per role and per round. The roles are plotted over the model (Figure 1) to show how players and their actions are positioned in respect of the system structure. During the game, the scope of all the roles remain the same, but certain roles can mature (CERT, DevOps, SecEng) by allocating resources to investments.

Table 1. Game data collected (per role and game round) plotted over the aggregated model structure.

Overview of Actions and Game Data Gathered for Each Game Role		
Role	**Actions**	**Game Data Gathered**
Attacker	Attacking and attack preparation	Strength of attack Security incident was handled in the war room (1) or not (0) Security incident resulted in damaging the defender (1) or not (0)
	Collecting revenue	The amount of damage paid for by the defender to the attacker
CERT	Resource allocation	Number of CERT resources in the game
	Proactively analysing threat intelligence or improving responsive policies	Number of CERT resources allocated to improvements.
SecEng	Resource allocation	Number of SecEng resources in the game
	Selecting, investing, maintaining and decommissioning security capabilities	Total level of capabilities in the data centre (=indication of maturity and number of capabilities). Data centre can handle up to 5 different capabilities (of 9 available)

Table 1. *Cont.*

Overview of Actions and Game Data Gathered for Each Game Role		
Role	**Actions**	**Game Data Gathered**
DevOps	Resource allocation	Number of CERT resources in the game
	Improving responsive policies	Number of DevOps resources allocated to improvements
	Maintaining business operations (customer transactions)	Number of DevOps resources allocated to serve customers
CEO	Gaining and maintaining customers	Number of customers
	Reporting on cyber-security incidents	Cyber-security incident was handled in the war room (1) or not (0) Cyber-security incident resulted in damaging the defender (1) or not (0)
	Financial performance	Total income, total resource costs and net result Amount allocated to operational reserve and to strategic reserve Income loss (amount of missed income caused by lack of DevOps resources serving customers) Supervisor warning (given when net result is negative)
	Resource allocation	Number of CEO resources (these are fixed in the game).

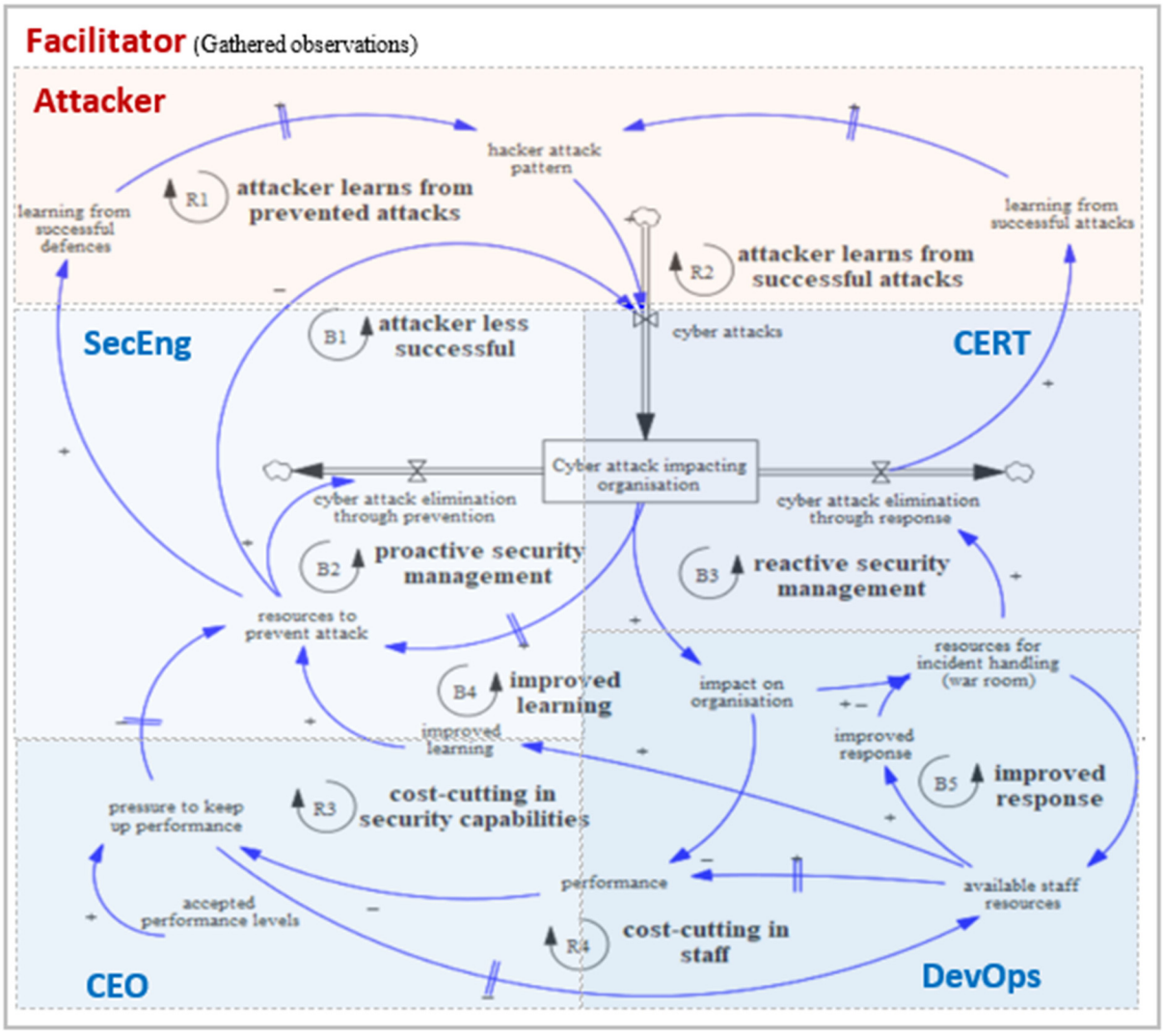

Figure 2. Game roles plotted over the aggregated model structure.

3.4.2. Round Sequence

Each game round consists of four phases.

Phase 1: Budget allocation. In this phase, the CEO allocates a budget to all other team members based on their requests and the available funds. This budget provides the means and boundaries for all blue team members' actions in the game.

Phase 2: Resource allocation. Here, the available budget can be used by the blue team managers. They can hire or fire human resources and allocate them across business, security or improvement processes. The improvement process costs resources but can generate future advantages in the game.

Phase 3: Attack processing. In this phase, incoming attacks need to be resolved by preventive measures laid out by the Security Engineer or CERT Manager and by responsive actions by the CERT Manager and DevOps Manager. If attacks are not resolved entirely, the resulting damage needs to be paid for by the CEO.

Specific to the DevOps manager, the resources needed for responsive actions lower the resources available for business operations. If the available resources for business operations are less than what is required to serve the customers, the revenue generation is impacted, reducing the income generated (income loss).

Phase 4: Financial results. The decisions taken during phases two and three will impact the cost and income generation during the game. Resource maintenance and the damage caused by successful attacks increase costs. On the other hand, sufficiently staffed business operations ensure that customers are being served, which will generate income. In this phase, the CEO calculates the costs paid and income received. The CEO allocates any remaining available funds to the strategic reserve (these funds cannot be used in the game itself anymore) or saves the money to cope with future losses and unexpected costs.

3.4.3. Game Board

The game board (Figure 3) is a pictorial representation of the underlying SD model, considering the game setting and sequence. To support the players' decisions, the game board also contains guidance stating the rules and explaining the game turn sequence.

3.4.4. Winning Criteria

The team with the highest strategic reserve at the end of two hours playing with a maximum of 20 turns wins the game, while the team with a strategic reserve below zero loses the game. The game uses a time limit instead of a fixed number of turns to simulate market forces and put pressure on the defenders' decision-making process. In addition, a team receiving more than four warnings, which can occur each turn during which a team incurs more costs than revenues (losses), also loses the game. These warnings, called supervisor warnings, mimic the real public regulator institution aiming at maintaining stability and prudence in the financial system. The warning indicates the presence of an emerging and risky financial situation in the game which requires improvement. The facilitator issues these warnings in the game.

3.4.5. Gameplay Testing and Validation

The challenge of translating the real-world situation to a playable game setting shaped the design of the game and its gameplay. The board game was tested through several playtests in which participants' feedback was collected and used to improve the board game for the next playtest. The iteration lasted until no significant concerns were raised anymore. Playtests are important because the game board design and game instructions must be clear, and the artificial organisation and game dynamics must be realistic. We conducted a total of four playtest sessions with different experts in cyber-security. Two people involved also participated in the modelling construction sessions, while the other two were chosen from a randomly selected group of players with backgrounds in IT, security, operational risk management or team management from a Fortune 500 organisation in the financial industry.

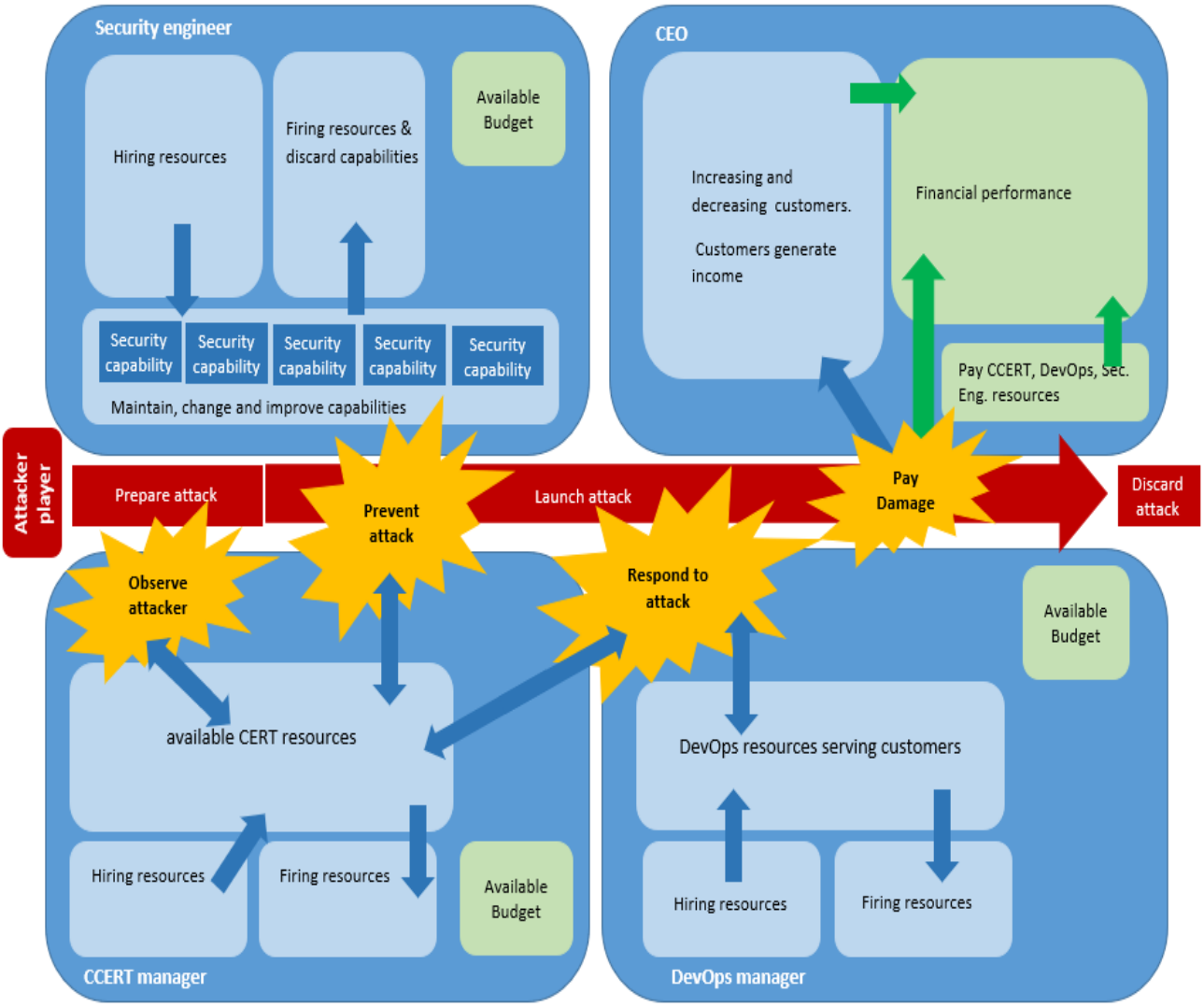

Figure 3. Simplified game board overview. The complete gameboard is visible in Appendix B.

The randomly selected group of participants all confirmed the underlying structure, outputs and the realism of the emerging dynamics. Their reactions were:

"This budget process is like our organisation. We have the same struggle."

"Yes, these are actually the dependencies I observe within our organisation."

"This is very realistic! We should play this with the people in my network."

This contributed to increasing the confidence in the reliability of the model and game structure. In addition, this feedback showed that two (partially) scripted rounds were sufficient to familiarise players with the game's rules. Based on the feedback received during these tests, the game board was improved with a "game turn" table (visible in Appendix B), and the concept of storytelling was used for the attacks. During the game, each launched attack took the form of a small, scripted story that covered the attack that occurred, the security capabilities that could defend against the attack and the impact of the attack (called damage).

4. Results and Analysis

We analysed the results of the gameplay rounds with the aim of increasing our understanding of strategic cyber-security decision-making. Our analysis was based on the data gathered from playing 16 games with 132 voluntary participants (as well as several facilitators who were guiding the games). A total of 240 game rounds were played over the course

of these 16 games. There were seven winning teams who played a total of 116 rounds and nine losing teams who played a total of 124 rounds. In our analysis, winning and losing teams' results were grouped following the winning criteria.

The domain knowledge and competences among the participants (all of whom had at least five years of experience) were distributed as follows: of the 132 participants, 73 were cyber-security specialists, 17 were senior managers, and the others held various roles related to overall organisational security. Game data (i.e., the values over time of the bank finances and the available human resources) were recorded at the end of each round, and the facilitators monitored the data-capturing process to ensure that no error was made in this process. The facilitators also took notes of any relevant behaviours and discussions taking place among the players in the teams they supervised. At the end of each game session, the player playing the CEO role was asked to anonymously fill out a questionnaire (see Appendix A) about his or her perception of their team's (blue team) performance. Specifically, the questionnaire asked about strategic priorities regarding the prevention of incidents, the response to incidents, the capability to serve customers, the policy for lowering costs and how well the team executed their strategies. This served as a participant's self-evaluation regarding their decisions and applied strategy. About half of the CEO players filled out this questionnaire.

4.1. Empirical Evidence from Gameplay and Questionnaires

We analysed the differences between winning and losing teams in terms of their decisions concerning: aggressiveness in the marketspace, resource allocation (e.g., staff members), security strategy and financial performance over time. Each of these will be explained in this section.

Firstly, aggressiveness in the marketspace was measured by the number of rounds played within a two-hour timeframe. The seven winning teams played, on average, 17 rounds, with a range between 14 and 20 rounds. The losing teams played, on average, 14 rounds, with a range between 7 and 16 rounds. According to the Mann–Whitney "U-test", there was no significant difference between the number of played rounds for the winning and losing teams.

Secondly, resource allocation was measured by the number of staff allocated to the CERT and DevOps departments. We focused only on these resources because they showed the most distinctive behaviour over time. They were plotted for each round, thus allowing us to compare time series for winning and losing teams (Figure 4). From the figure, it is possible to note how the two groups differed in staffing policies: winning-teams displayed a more stable staffing policy, whereas losing teams showed an unstable one, characterised by sudden increases and reductions in the staff volumes. On average, the losing teams had approximately twice as many human resources compared to the winning teams.

Due to the time delays in the hiring and firing process caused by the onboarding process and dismissal process, respectively, human resources temporarily generate costs while they are not directly productive to the department. In other words, hiring and firing incurs costs to the organisation. To analyse this aspect further, we looked in more detail into the headcount fluctuation for the DevOps and CERT resource levels. For each round, we determined the difference compared to the previous round and calculated the absolute difference between these numbers: this is what we refer to as the "fluctuation effect" (FE). We calculated the FE means and standard deviation for both winning and losing teams per round. High levels of fluctuations (represented by high values of FE) indicate that a high number of staff resources are in the process of being hired or fired, instead of being productive for the defender. If a player alternates between hiring and firing staff resources more often or for greater numbers of staff, this is reflected in a higher FE. Table 2, below, reports these results. The losing teams had a 233% higher average fluctuation effect compared to the winning teams and a significant difference in hiring and firing CERT and DevOps resources compared to the winning teams. This means the losing teams had relative low levels of productive resources, because most of their headcount was

in the process of being hired or fired and therefore was not employed in operations (hence was not productive).

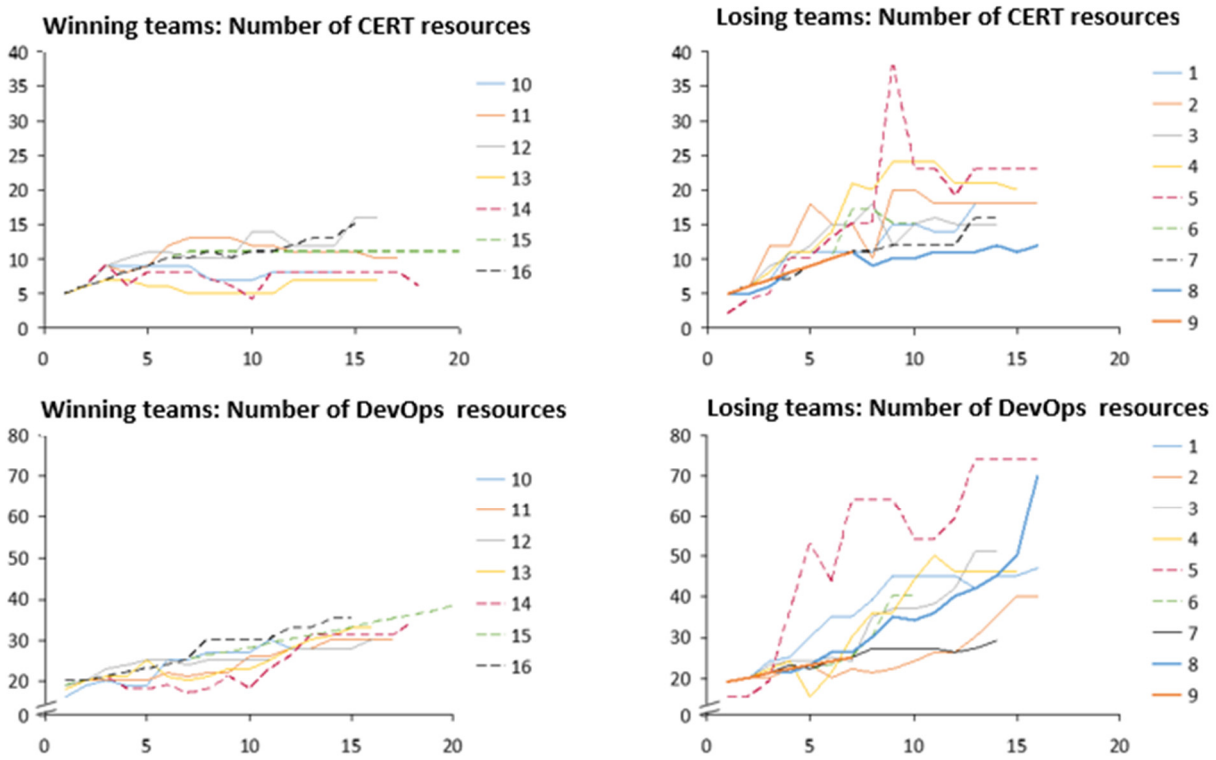

Figure 4. Difference in CERT and DevOps staff for winning and losing teams over time (per round). Each curve represents a team playing the game.

Table 2. Average and standard deviation (SD) of the fluctuation effect (FE) for CERT and DevOps resources for the winning and losing teams and *t*-test statistics. *** = 0.01 significance.

	Information	## Rounds	CERT ***	DevOps ***
Game data	Losing teams-FE average	124	1.6	2.8
	losing teams-FE SD	124	2.9	4.1
	Winning teams-FE average	116	0.7	1.2
	Winning teams-FE SD	116	1.0	1.3
***t*-test statistics**	difference		0.93	1.62
	standard error		0.285	0.396
	t-statistics		3.262	4.093
	CI		95%	95%
	DF		238	238
	significance level		0.0013	0.0001

We also performed a regression analysis to see if the difference in FE between losing and winning teams could be explained by the overall strength of the incurred cyber-attacks or side-effects. The equation of the regression analysis was $Y = \beta 0 + \beta 1X1 + \beta 2X2 + \epsilon$, where Y = occasions of FE (we accounted for such an effect when its value was equal to 2 or higher); X1 = occurrences of income losses; and X2 = strength of the attack. Based on the regression analysis outcome, the losing teams' sudden increases in resource hiring/firing could be related to a specific side-effect, the occurrence of income loss ($F = 22.6$, $p < 0.01$ ***). This suggested that losing teams' resource allocation tactics were reactively driven by business disruptions.

Thirdly, we investigated the adopted cyber-security strategies for both the winning and losing teams. This was a comparative analysis of the defenders' reactions to attacks

in relation to capability maturity and the attacks' strength. To do so, we separated the defender actions into the following categories: 'prevent' (resources are allocated to increase attack prevention), 'response' (resources are assigned for resolving an incident that has occurred) and 'pay damage' (the CEO has to pay for the damage caused by the attack). We grouped the winning and losing teams, averaged the total strength of the attacks incurred by the teams per round and plotted the obtained results of attack strength over time against defender actions. The strength of the attack launched by the attacker in each round was the same for all the teams. This is visible in the attacker strength curve. The results are shown in Figure 5, where the total combined results of the defender (combination of prevent, response and pay damage) and the attack strength curves are reported on a relative scale of 0.0 to 1.0 (left axis—1.0 corresponds with 100% of the launched attacks). According to the Mann–Whitney U test, there was a significant difference ($z = -2.652$ and $p = 0.00402$) between losing (prevented 50% of all attacks launched during the played rounds) and winning teams (prevented 63% all attacks launched during the played rounds): the latter prevented significantly more attacks (about 13 percentage points).

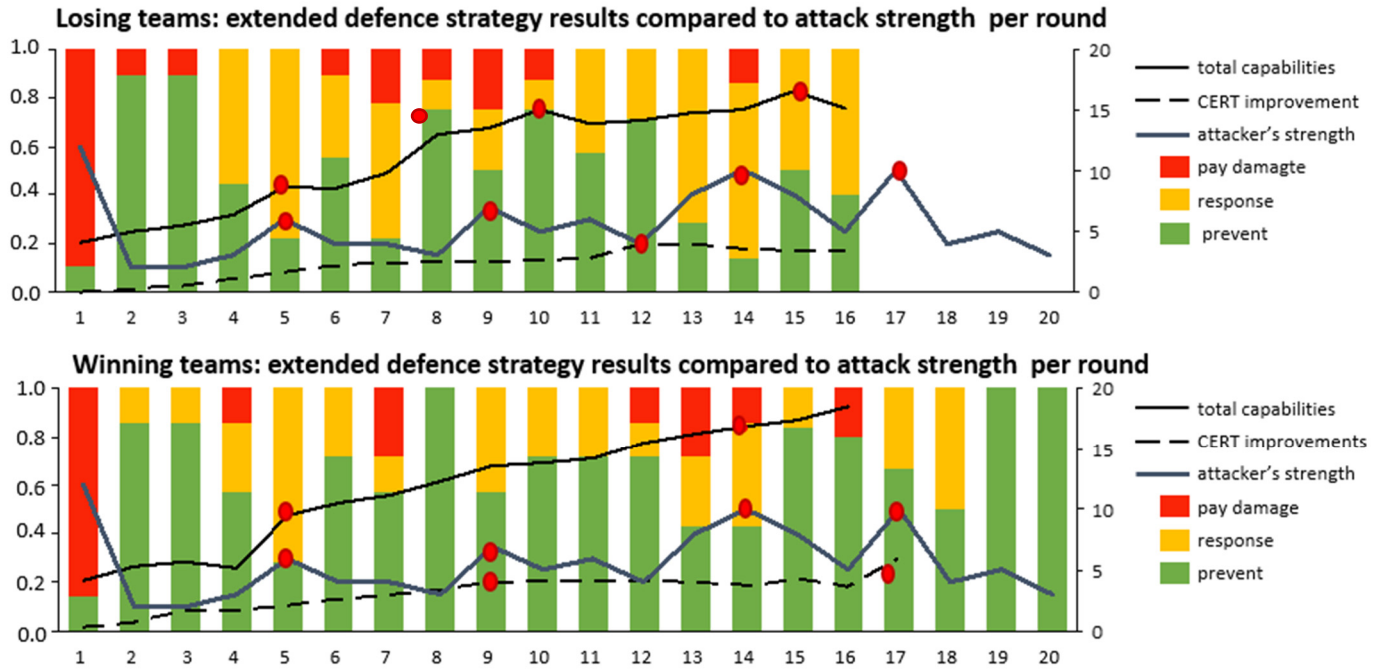

Figure 5. The average percentage of cyber-security strategies used by teams that prevented an attack, responded to an attack or paid for damage caused by an attack per round for the losing and winning teams is plotted against the attack strength of the launched attack, the cumulative average number of CERT improvements and the cumulative average total security capabilities per round. Red dots mark game-changing events.

To further investigate why the winning teams prevented more attacks, we compared the attacks' strength, the development of the defences and the CERT improvement over the game for both groups. The value of all these variables could be between 0 and 20 (Figure 5—right axis). The total (defence) capabilities curve resembled the maturity of the defences: the defender's ability to cope with attacks. The higher this number, the more able the defender was to handle diverse and powerful attacks. For example, a low defense capability level (value 1) was effective against a weak attack, while a mature one (value 2 to 6) was able to mitigate stronger attacks. Multiple defence capabilities are needed against different attack forms. The average value was calculated for both the winning and losing teams. The number of CERT improvements resembled aspects of the resilience organisation through the implementation of different sophisticated learning capabilities. A higher number represented a more resilient organisation. The average

value was calculated for both the winning teams and the losing teams. In the game, CERT improvements provided the benefits of the defender's early anticipation of attacker behaviour or of more effective responses by CERT. Examples of these improvements are: advanced threat intelligence, informing the defender about upcoming attacks; time-delay reduction in response processes due to runbooks; and higher productivity, due to improved incident response processes.

Crucial events and changes in the development of these metrics in Figure 5 are marked with a red dot. From the analysis of the plotted data (Figure 5), it was possible to note a particular sequence of strong attacks in rounds 5, 14 and 17. Interestingly, winning teams increased their defensive capabilities (in rounds 5 and 14) or improved the effectiveness of their response to incidents (CERT function, in round 9) in the same round of the attack. This meant that they started the process of defence and response capability improvement before the attacks started. On the contrary, the losing teams realised that they would need the very same improvements only several rounds after the attacks had occurred effectively (round 7, 12 and 15, respectively). This suggests that, in contrast to the winning teams, the losing teams did not anticipate the attacks but were more inclined to react only after they perceived that successful attacks were starting to occur.

Fourthly, the ex post CEO questionnaires were compared with the game results to try to assess the players' self-evaluations of their performances. The questionnaire collected information on the importance of specific strategic priorities during the game (on a scale of very important (1) to not important (5)) and on how well they perceived their strategy execution to be in respect to these priorities (ranging from very well (1) to very poor (5)). We analysed the players' perceived priority of financial strategy performance and execution and compared this with the actual financial results of the game. Table 3, below, reports the results of this comparison. On the right-hand side, the results of the questionnaire are visible. These questionnaires did not indicate any significant difference between the winning and losing teams in terms of perception. On the left-hand side, the results of the game data are visible. Clearly, they show a significant difference in performance between the winning and losing teams for all the criteria for winning the game or for being knocked out of the game. Interestingly, although the game data show a significant difference between the winning and losing teams, this distinction cannot be observed in the CEOs' answers. The questionnaire indicated that the perceptions of execution and importance were slightly higher for the winning than the losing teams, but not significantly so. This suggested that the losing teams might not have been aware of their poor decision-making. This was especially interesting because the losing teams also received multiple warnings about their financial performance (during every round in which the team had a negative balance, i.e., when costs were higher than income, the facilitator delivered a warning to the team).

Table 3. Game results versus questionnaire: financial performance and *t*-test statistics. *** = 0.01 significance.

				Game Results-Financial Performance					**Questionnaire-Financial Performance**	
	Team	**##**		**cumulative Supervisor Warnings *****	**Cumulative Net Result *****	**Cumulative Operational Reserve *****	**Cumulative Strategic Reserve *****	**##**	**Execute**	**Importance**
Game data	winning teams	7	average	3.0	109.4	56.1	78.4	4	3.0	2.3
			SD	1.7	67.3	42.8	45.8		0.8	1.0
	losing teams	9	Average	8.0	−69.8	−94.2	13.9	5	3.6	3.8
			SD	3.8	110.8	199.5	16.2		1.7	1.5
***t*-test statistics**		difference		5	−179.2	−150.3	−64.5		0.6	1.45
		standard error		1.552	76.693	77.3	16.3		0.931	0.878
		T-statistic		3.221	−3.757	−1.944	−3.952		0.645	1.651
		CI		0.95	0.95	0.95	0.95		0.95	0.95
		DF		14	14	14	14		7	7
		significance level		0.0062	0.0021	0.0722	0.0014		0.5398	0.1427

4.2. Game Observation Results

The purpose of this section is to discuss if the observations made during gameplay indicated distinctive behavioural differences between the winning and losing teams. Expressions of feelings, emotions and observed behaviour are signals of how people appreciate the multiple feedback loops, time delays and nonlinearities in the dynamic complex problem at hand [18].

The trained facilitators made observations during the game about important game dynamics at their own discretion. As a result, 40 observations were gathered, consisting of quotes (14) and observed behaviours (22) during the game and feedback about the game (4). The observations collected by the facilitators were then analysed following inductive thematic analysis.

Ten of these observations referred to winning teams. Nine of them were about observed behaviour. The nature of these observations indicated that the winning teams showed more constructive and cooperative behaviour, being calm and prudent and seeking justification for their decisions. This created a context in which individual knowledge was actively shared with peer decision-makers, in line with Desouza's findings [75]. We include one quote from a winning team member:

DevOps manager to the team, regarding the decision: *"We do it for our customers."*

A total of 30 observations referred to losing teams. Fifteen observations were about behaviour, twelve were quotes and three were feedback about the game. Losing teams showed less constructive decision-making behaviour in 14 out of 15 observations: this was seen through the presence of dominant decision-makers and specific behaviours exhibited by decision-makers such being self-interested, covering up for poor decisions, and fearing the consequences of decisions. Ten out of twelve quotes related to blaming others, frustration, helplessness and observations about wrong decisions. The following are some quotes from members of the losing teams:

CERT manager to CEO: *"Are we making money by the way?"* CEO responded: *"I have no idea what you did at all."* (This quote suggest that the CERT manager would have liked to have had feedback on their team performance, and that the CEO did not understand what the performance drivers within his organisation were.)

CEO: *"You invested in the wrong capabilities"*. Security Engineer responded: *"Thanks for approving all my budget. Why can't I double my request?"* (This quote suggests that the CEO disagreed with the Security Engineer's decisions, while the Security Engineer had an opinion regarding his lack of budget.)

Blue team: *"Keep the hack quiet or we blame the CEO. We outsource the hack."* (This is an example of trying to blame others.)

"I am not going to pay." said an exasperated CEO when the attack was not mitigated (however, it was too late, since paying the damage of an attack was necessary when an attack was not mitigated).

"Why are these attacks not blocked by our defences?" (The blue team was so focused on the game board that they did not notice the attacker was standing next to them looking over their shoulder to their defences.)

Although research indicates that negative financial consequences contribute to better security policy compliance [2], our game results may indicate otherwise. Negative financial consequences in our game occur when the running costs are higher than income. They may evoke an overall loss that triggers supervisory warnings and potentially leads to losing the game. Although the nine losing teams together received a total of 72 warnings that they were making a loss, only a few corrective actions were observed. In fact, only some of the losing teams proactively addressed the high numbers of wasted resources. Some of the observed participants' remarks were:

DevOps: "We have got more Security Engineers than DevOps Engineers";

DevOps: "We need more, too much here, too much there. This is killing us";

CEO: "Are you sure, 24 CERT resources? 24?", CERT manager: "You are the CEO!"

5. Discussion

The data obtained in the game sessions provided valuable insights into the relationship between cyber-security investments and decision-making processes. Table 4 summarises and compares the results obtained for the winning and losing teams. The significant difference in game performance between the winning and losing teams can be traced back to their decision-making. The winning teams focused on prevention, and their higher CERT capabilities, developed earlier in the game, contributed to their organisation's resilience to cyber-attacks. Such improvements provided the means to anticipate upcoming attacks earlier (due to threat intelligence, whereby these threats became known to the defender) and to respond better (because of runbooks with improved and updated response procedures). While the winning teams proactively invested in prevention strategies sufficiently far in advance, the losing teams, on the contrary, manifested 'reactive' behaviours (i.e., taking action after incidents occurred), resulting in less effective defence strategies, higher human resource fluctuations and, ultimately, poorer financial performance. Taken at face value, this outcome may not seem surprising, as it is common knowledge that "prevention is better than cure". However, what appears staggering is that 9 out of 16 teams consisting of recognised experts did not opt for what is generally recognised as the most convenient and safest approach. We specifically attribute this fallacious behaviour in the losing teams' decision-making process to the use of heuristics, which led to the wrong decisions. From the game results, we could infer that the difference in game performance between the losing and winning teams came down to how information about performance acquired by the players during the game was digested individually and evaluated collectively [76], plus how they approached and anticipated the uncertain future rounds as a team. While the winning teams were able to anticipate that attacks could grow much worse in future rounds if they did not invest in defence and CERT capabilities, we inferred that the losing teams based their decisions on the initial (apparent) relative ease with which they dealt with the first attacks (the 'past'). The losing teams' approach may be related to decision-making under complex and uncertain conditions that evoke the use of heuristics [17,33,41]. Several types of heuristics in which past and initial information heavily affect judgements can be recognized. For example, anchoring and adjustment (relying on the first pieces of information received, the 'anchor' [77]), familiarity (the past system behaviour solutions hold true also for the present [78]), and educated guesses (decisions based on the individual previously stored information [31]) could be the mechanisms responsible for this.

The insights gained on the basis of the observation of the participants' behaviour can assist in identifying other potentially harmful decision mechanisms that contributed to the teams' losses. For example, the losing teams' decision-making processes seemed less constructive compared to the winning teams'. In terms of directly observed behaviour, signs of frustration, blaming, helplessness and non-cooperative and nonconstructive collaboration emerged. Moreover, the players in the losing teams did not always consider the effects of their decisions on the other actors in the game. In this context, the signs of frustration and agitation may have signalled the use of emotion-related heuristics [43]. In particular, affect (emotions and feelings related to a stimulus leading the decision process [79]) and social proof (social context—mimicking and approval—significantly determining decisions [80]) seem to be likely candidate heuristics underlying the observed behaviour.

Concerning the development of staff resources over time in the losing teams (see Figure 4), an interesting trend emerged. In fact, not only did the losing teams have, on average, about twice as many resources compared to the winning teams, but their trends showed concerning oscillations, in which the amplitude of the peaks increased over time. We advance two connected reasons to explain why this happened.

Table 4. Summarized results specified for losing and winning teams.

Area	Topics	Source	Losing Teams	Winning Teams
Game performance	Resource allocation strategy	Game results	• Incidents drive hiring and firing decisions of staff • Inefficient resource utilisation	• More sustainable staff development • Efficient resource utilisation
	Security strategy		• Reactive approach and learning from past incidents • Prevent 13 percentage points fewer attacks	• Focus on prevention and resilience • Prevent 13 percentage points more attacks
	Financial performance		• Making losses • Low strategic reserves	• Making profits • Adding value (high strategic reserves)
Decision-making processes	Self-evaluation	Questionnaire	Acceptable performance	Acceptable performance
	Decision-making environment	Facilitator observations	Dominant decision-makers, being self-interested, covering up, and fearing consequences	Constructive and cooperative behaviour, being calm and prudent, and seeking justification for decisions
	Expressions		Blaming others, frustration, helplessness, ignoring criticism on decisions	None

Firstly, when the losing teams were strongly hit by the attacks, they tended to overreact. Namely, they allocated too many resources and hired too may staff to deal with the response strategy (DevOps and CERT department), in order to mitigate the consequences of the incidents. However, these staffing policies were not remunerative and contributed to a decrease in income, so these teams tended to fire the resources they believed to be in excess in the following rounds, until they were hit again by an attack and hired even more staff for the response strategy, starting the oscillation again. This decision process might be partially explained by representativeness heuristics [77], which describes a situation in which a decision is not necessarily taken based on the entire past experience available to the decision-maker and its 'averaging', but is driven by what the mind believes to be the most salient and intense moment in its experience (i.e., the occurrence of a strong attack).

Secondly, in addition to this, a systemic element contributed to the emergence of the oscillating trend. Hiring and firing resources in this game takes time (measured in rounds) and costs money, and the staff in the process of being hired or fired cannot be actively used in the game. This represents a typical system delay [17]. Therefore, the results of these actions need some time to manifest themselves, while in the meantime, the system condition may have changed, again requiring a change in the course of action. More specifically, when a severe incident occurred, the teams rushed to increase their response staff (overreaction), and once the staff was hired, it was later perceived as unnecessary (due to changed conditions), and they were subsequently laid off (leaving the team again unprepared for future strong attacks). Interestingly, the losing teams did not put in place any corrective (balancing) structural mechanisms that could help them limit their resources growth. Furthermore, these dynamics resembled those of an organisation caught in the previously evoked "capability trap" [35]: since resources have not been allocated on time to prevention activities, capital needs to be allocated to response strategies so as to mitigate the consequences instead of making up for the prevention improvements that were skipped earlier.

Finally, the questionnaire results showed no significant difference in perceived and actual game performance between the winning and losing teams. This result supports the assumption that the losing teams were not aware of the consequences of their decisions that resulted in their poor performance.

6. Limitations and Future Research

The different usages of heuristics appear to be the most convincing reasons for the diverse decision paths adopted by the winning and losing teams. However, other explanations, which we did not observe in our research, could be related to the effects of conflicting attitudes and beliefs in the teams [81], learned helplessness making decision-makers not able to choose between the possible alternative interventions [82], or the willingness to experiment with parameters by participants [83]. The "Red versus Blue" cyber-security board game also aimed to improve the decision-making of the participants by encouraging them to learn from their experiences in complex situations. In this way, the game provides a similar experience to simulation-based training [20] and gaining familiarity [84], while the game pays extra attention to human interactions. However, to support this claim, further research in this direction needs to be carried out. Measurements of the impact of our game on participants' awareness and system understanding through both ex ante and ex post surveys could further increase our comprehension of the game's effects on the understanding of interdependencies and resulting dynamics in cyber-security. Additionally, video observations of the gaming sessions may provide more quotes and observations about participant behaviour. Hence, we believe that the game should continue to be improved, refined, and tested in other contexts and situations in order to further assess its effectiveness as a learning tool and as an experimental teaching method. In this respect, it is important to remember that learning processes differ from person to person and depend on a willingness to perform demanding cognitive activities [85]; hence, the players' motivation may have affected the outcomes of the exercise.

Concerning strategic cyber-security decision-making, our hypotheses regarding the role of heuristics need further validation through experimental settings and the direct observation of real cases (despite the challenges). More generally, and in line with Jalal et al. [41], further research and empirical tests should be conducted on the behavioural biases about the complexities of decision-making in the field of cyber-security economics, as we believe that there is still a lot of untapped potential.

7. Conclusions

In this paper, we reported on our experience of developing a serious gaming approach to mimic real-life strategic cyber-security decision-making and using it to observe cyber-security decision-making practices. Gaining a better understanding of these topics is important, since people experience great difficulties understanding the effects of multiple feedback loops, time delays and accumulations in complex dynamic systems [17], such as cyber-security management, and thus they may employ simplified decision approaches (heuristics) to lower the cognitive burden and make potentially fallacious decisions [38–40]. Specifically, to construct the game, we firstly built an SD model capturing real-life strategic cyber-security decision-making and validated it, and then a board game was developed based on this formal model. During the evaluation phases, the game was validated and received positive feedback from several experts in the field pertaining to its realism. This suggests that the game is a promising tool not only for experiments concerning cyber-security decision-making but also for training participants. In fact, a serious gaming environment allows the participants to explore, experience and learn from their actions without impacting the real business environment. As was previously stated, 9 out of 16 playing teams lost since they had a poor financial performance. Compared to the winning teams, they had fewer security capabilities in place and showed mostly reactive behaviour and unstable staffing policies (continuous alternation of hiring and firing). The game's experimental setting allowed us to look directly at the decision-making process

that led to the lack of success. In this process, our observations suggested the use of heuristics, connected to the incorrect interpretation of the representativeness of past events, which led the losing team to overreaction (amplified hiring and firing in response to income loss) and reactive strategies (increasing defensive capabilities only after being impacted by the attack—hence displaying a lack of preventive vision). We assumed that the critical attitudes towards other team members were displayed as a response to the loss of control over the deteriorating situation, as a way for players to exonerate themselves and identify the problems in other bank sectors. Such emotivity could have been a signal that affect [79] or social proof [80] heuristics were also at play, namely that peers' judgments could have significantly influenced decision-makers' behaviours. Most importantly, from the comparison between the game results and the perceptions measured via the ex post questionnaires, we observed that the players in the losing teams were not aware of their insufficient individual performance. These decision-making strategies and heuristics that emerged in the game could be helpful to explain specific organisations' cyber-security incidents caused by suboptimal (or even poor) decision-making behaviours.

In conclusion, there are three important practical implications:

Firstly, the interconnectedness of cyber-security with business operations, IT, finance and risk management advocates an integrated and holistic decision-making approach.

Secondly, without holistic support, cyber-security managers may be overwhelmed by the complexity of the system in which they operate; a good understanding of the underlying systemic structure may be a partial antidote to the use of fallacious decision strategies. In addition to this, future work should attempt to develop instruments and methods that minimise the avoidable biases related to the use of heuristics in cyber-security management. An improvement to the tools available to counter heuristics in cyber-security management would lead to more secure organisations experiencing lower cyber incidents.

Thirdly, the game results showed that organisations prioritising prevention and response capabilities following the principles of a resilient organisation outperformed those who were not proactive and waited for an incident before acting, confirming existing knowledge [13].

Overall, these insights are expected to enhance the comprehension of decision-making processes in the context of cyber-security financial management in organisations. Any progress on this aspect will help to avoid resource misplacement, decrease cyber vulnerability and support cyber-security specialists in improving cyber-defences in their organisations [59]. Finally, this work further supports the evidence of the relevance of using serious gaming instruments, in particular SD-based instruments [48], for improving cyber-security management [75].

Author Contributions: Conceptualization, S.Z., E.A.J.A.R. and M.v.K.; methodology, S.Z.; software, S.Z.; validation, S.Z., E.A.J.A.R. and M.v.K.; formal analysis, S.Z.; investigation, S.Z.; resources, S.Z., G.C. and S.A.; data curation, S.Z.; writing—original draft preparation, S.Z., G.C. and S.A.; writing—review and editing, S.Z., G.C., S.A., E.A.J.A.R. and M.v.K.; visualization, S.Z.; supervision, S.Z., E.A.J.A.R., M.v.K. and S.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not Applicable.

Informed Consent Statement: All participants joint these sessions voluntarily. Prior publication we received approval in written from appropriate management of target organizations.

Data Availability Statement: The data are not publicly available.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Self-Evaluation Questionnaire

Team number:

Please indicate what level and capabilities you have in your organisation					
Capability	Level 1	Level 2	Level 3	Level 4	Level 5
Security Event Monitoring					
Identity & Access MGT					
Vulnerability Management					
DDOS Protection					
Malware Protection					
E-Banking Fraud Protection					
Secure Prog. & Testing					
Key Management					
Awareness					
Please indicate what possible strategies are important to you and how well they were executed by your team					
Strategic Importance	1 Very Important	2	3 Average	4	5 Not Important
Prevent Cyber Attacks					
Respond to Cyber Attacks					
Serve Customers					
Lowest Costs					
Strategy Execution	1 Very Well Executed	2	3 Average	4	5 Very Poor Executed
Prevent Cyber Attacks					
Respond to Cyber Attacks					
Serve Customers					
Lowest Costs					
What went well and why:					
What went wrong and why:					
Number of intervention cards received from facilitator:					

Appendix B. Gameboard

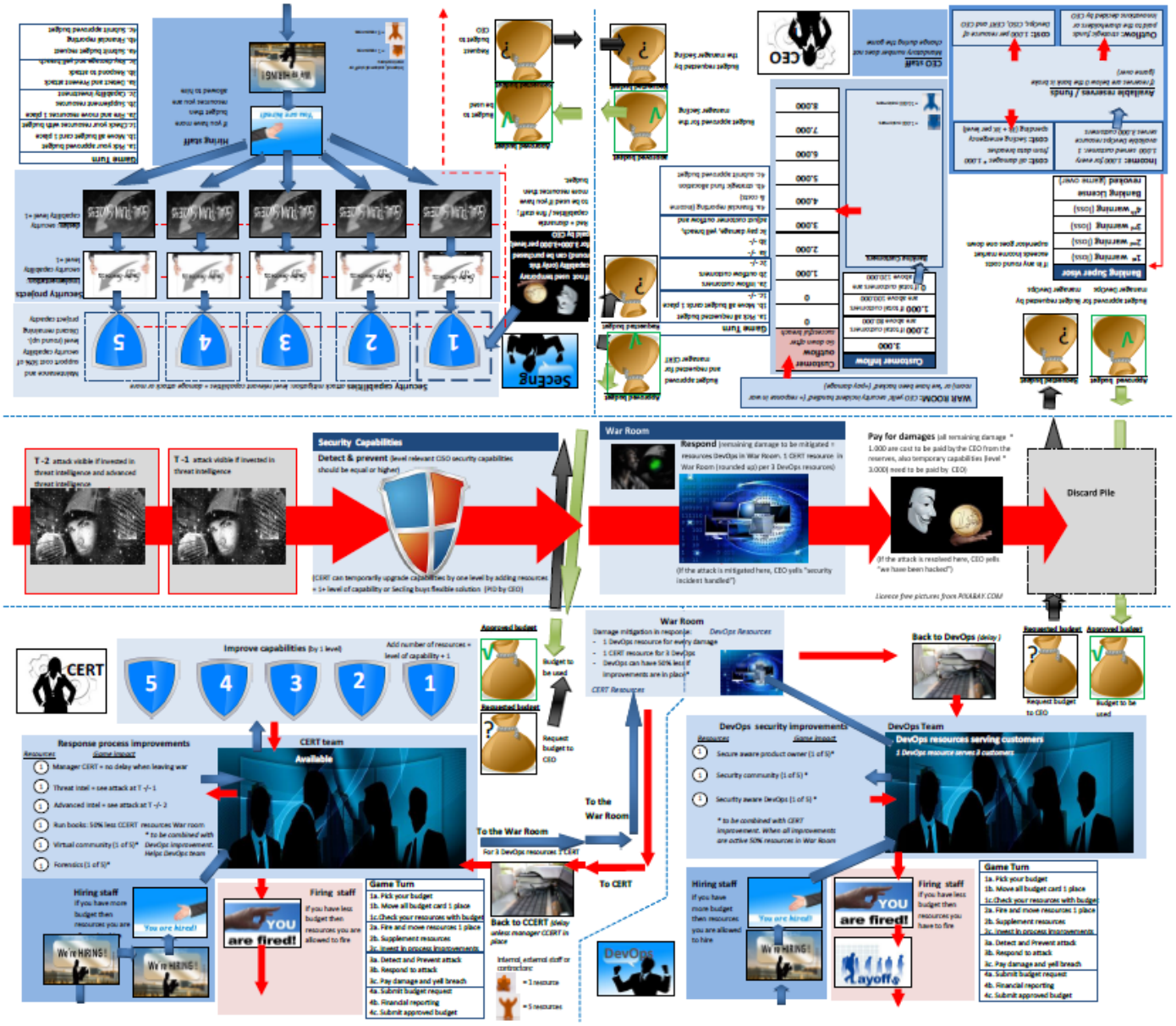

Figure A1. Game board overview.

References

1. Christina, Y.; Jeonga, C.Y.; Leeb, S.T.; Lim, J. Information security breaches and IT security investments: Impacts on competitors. *Inf. Manag.* **2019**, *56*, 681–695.
2. Goel, S.; Williams, K.J.; Huang, J.; Warkentin, M. Can financial incentives help with the struggle for security policy compliance? *Inf. Manag.* **2021**, *58*, 103447. [CrossRef]
3. Ritcher, F. 200,000+ Systems Affected by WannaCry Ransom Attack. Available online: https://www.statista.com/chart/9399/wannacry-cyber-attack-in-numbers (accessed on 21 December 2021).
4. GReAT. The Great Bank Robbery: The Carbanak APT. Available online: https://securelist.com/the-great-bank-robbery-the-carbanak-apt/68732/ (accessed on 21 December 2021).
5. Modderkolk, H. *Het Is Oorlog Maar Niemand Die Het Ziet*; Uitgeverij Podium: Amsterdam, The Netherlands, 2019.
6. Armenia, S.; Angelini, M.; Nonino, F.; Palombi, G.; Schlitzer, M.F. A dynamic simulation approach to support the evaluation of cyber risks and security investments in SMEs. *Decis. Support Syst.* **2021**, *147*, 113580. [CrossRef]
7. Goel, S.; Shawky, H.A. Estimating the market impact of security breach announcements on firm values. *Inf. Manag.* **2009**, *46*, 404–410. [CrossRef]

8. Zeijlemaker, S. Exploring the dynamic complexity of the cyber-security economic equilibrium. In Proceedings of the 34th International Conference of the System Dynamics Society, Delft, The Netherlands, 17–21 July 2016.
9. Zeijlemaker, S. *Cyber-Security Quantification: Founding a Structural Understanding of Its Dynamic Complexity*; Radboud University: Nijmegen, The Netherlands, 2017.
10. Zeijlemaker, S.; Uriega, J.D.; Kilanc, G.P. Malware dynamics: How to develop a successful anti-malware defence reference architecture policy. In Proceedings of the 36th International Conference of the System Dynamics Society, Reykjavik, Iceland, 6–10 August 2018.
11. Srinidhi, B.; Yan, J.; Tayi, G.K. Allocation of resources to cyber-security: The effect of misalignment of interest between managers and investors. *Decis. Support Syst.* **2015**, *75*, 49–62. [CrossRef]
12. Moore, T. The economics of cybersecurity: Principles and policy options. *Int. J. Crit. Infrastruct. Prot.* **2010**, *3*, 103–117. [CrossRef]
13. Zhou, Y.; Solak, S. Measuring and Optimizing Cybersecurity Investments: A Quantitative Portfolio Approach. In Proceedings of the 2014 Industrial and Systems Engineering Research Conference, Montréal, Canada, 31 May–3 June 2014.
14. Anderson, R.; Barton, C.; Böhme, R.; Clayton, R.; van Eeten, M.J.G.; Levi, M.; Moore, T.; Savage, S. Measuring the Cost of Cybercrime. In *The Economics of Information Security and Privacy*; Springer: Berlin/Heidelberg, Germany, 2013; pp. 265–300.
15. Schneier, B. CYA Security, Schneier on Security. Available online: https://www.schneier.com/blog/archives/2007/02/cya_security_1.html (accessed on 30 August 2015).
16. Tongia, R.; Kanika, J. Investing in Security—Do not rely on FUD. *Inf. Syst. Control. J.* **2003**, *5*. Available online: https://www.researchgate.net/profile/Rahul-Tongia/publication/238746543_Investing_in_Security-Do_Not_Rely_on_FUD/links/559e759508aea946c06a0880/Investing-in-Security-Do-Not-Rely-on-FUD.pdf (accessed on 21 December 2021).
17. Sterman, J. Modeling Managerial Behavior: Misperceptions of Feedback in a Dynamic Decision-making Experiment. *Manag. Sci.* **1989**, *35*, 321–339. [CrossRef]
18. Sterman, J. Teaching Takes off: Flight Simulators for Management Education "The Beer Game", October 1992. Available online: http://web.mit.edu/jsterman/www/SDG/beergame.html (accessed on 21 December 2021).
19. Vennix, J.A.M. *Group Model Building, Facilitating Team Learning Using System Dynamics*; John Wiley & Sons Ltd.: Hoboken, NJ, USA, 1996.
20. Sterman, J.D. Learning from Evidence in a Complex World. *Am. J. Public Health* **2006**, *96*, 505–514. [CrossRef]
21. Moore, T.; Duynes, S.; Chang, F.R. Identifying How Firms Manage Security Investment. In Proceedings of the Workshop on the Economics of Information Security (WEIS), Berkeley, CA, USA, 13–14 June 2016.
22. Hendrix, M.; Al-Sherbaz, A.; Bloom, V. Game Based Cyber-Security Training: Are Serious Games suitable for cyber-security training? *Int. J. Serious Games* **2016**, *3*, 53–61. [CrossRef]
23. Wouters, P.; van Nimwegen, C.; van Oostendorp, H.; van der Spek, E.D. A meta-analysis of the cognitive and motivational effects of serious games. *J. Educ. Psychol.* **2013**, *105*, 249–265. [CrossRef]
24. Tseng, S.S.; Yang, T.Y.; Wang, Y.J. Designing a Cyber-security Board Game Based on Design Thinking Approach. In *International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing. Advances in Intelligent Systems and Computing*; Barolli, L., Xhafa, F., Javaid, N., Enokido, T., Eds.; Springer: Cham, Switzerland, 2019; pp. 642–650. [CrossRef]
25. Olano, M.; Sherman, A.; Oliva, L.; Cox, R.; Firestone, D.; Kubik, O.; Patil, M.; Saymour, J.; Sohn, I.; Thomas, D. Security Empire: Development and Evaluation of a Digital Game to Promote Cyber-Security Education. In Proceedings of the 2014 USENIX Summit on Gaming, Games and Gamification in Security Education, San Diego, CA, USA, 18 August 2014.
26. Falco, G.; Eling, M.; Jablanski, D.; Miller, V.; Gordon, L.A.; Wang, S.F.; Schmit, J.; Thomas, R.; Elvedi, M.; Maillart, T.; et al. A Research Agenda for Cyber Risk and Cyber Insurance. In Proceedings of the 2019 Workshop on the Economics of Information Security, Boston, MA, USA, 3–4 June 2019.
27. Kenneally, E.; Randazzese, L.; Balenson, D. *Cyber Risk Economics Capability Gaps Research Strategy. United States Department of Homeland Security*; Science and Technology Directorate: Washington, DC, USA, 2018. [CrossRef]
28. Sterman, J. *Business Dynamics: Systems Thinking and Modeling for a Complex World*; Irwin/McGraw-Hill: Boston, MA, USA, 2000.
29. Neth, H.; Gigerenzer, G. Heuristics: Tools for an Uncertain World Emerging. In *Trends in the Social and Behavioral Sciences*; Scott, R., Kosslyn, S., Eds.; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2015; ISBN 978-1-118-90077-2.
30. Tversky, A.; Kahneman, D. Judgement under uncertainty: Heuristic and biases. *Or. Inst. Res. Bull.* **1973**, *13*, 1.
31. Kahneman, D.; Slovic, P.; Tversky, A. *Judgement under Uncertainty: Heuristics and Biases*; Cambridge University Press: Cambridge, UK, 1982; ISBN 9780511809477.
32. Myers, D.G. *Social Psychology*; McGraw-Hill Higher Education: New York, NY, USA, 2010.
33. Größler, A.; Bleijenberg, I.; Vennix, J. 10 Years on Average Doesn't Mean 10 Years in Any Case—An Experimental Investigation of People's Understanding of Fixed and Continuous Delays. In Proceedings of the International System Dynamics Conference, Washington, DC, USA, 24–28 July 2011.
34. Martinez-Moyano, I.J.; Morrison, D.; Sallach, D. Modeling Adversarial Dynamics. In Proceedings of the 2015 Winter Simulation Conference, Huntington Beach, CA, USA, 6–9 December 2015.
35. Repenning, N.P.; Sterman, J.D. Capability Traps and Self-Confirming Attribution Errors in the Dynamics of Process Improvement. *Adm. Sci. Q.* **2002**, *47*, 265–295. [CrossRef]
36. Hofstede, G. Management control of public and not-for-profit activities. *Account. Organ. Soc.* **1981**, *6*, 193–211. [CrossRef]
37. Kahneman, D.; Tversky, A. Prospect Theory: An Analysis of Decision Under Risk. *Econometrica* **1979**, *47*, 263–292. [CrossRef]
38. Rosoff, H.; Cui, J.; John, R.S. Heuristics and biases in cyber security dilemmas. *Environ. Syst. Decis.* **2013**, *33*, 517–529. [CrossRef]

39. Gomez, M.A.; Villar, E.B. Fear, Uncertainty, and Dread: Cognitive Heuristics and Cyber Threats. *Politics Gov.* **2018**, *6*, 61–72. [CrossRef]
40. Van Schaik, P.; Renaud, K.; Wilson, C.; Jansen, J.; Onibokun, J. Risk as affect: The affect heuristic in cybersecurity. *Comput. Secur.* **2020**, *90*, 101651. [CrossRef]
41. Jalali, M.S.; Siegel, M.; Madnick, S. Decision-making and Biases in Cyber-security Capability Development: Evidence from a Simulation Game Experiment. *J. Strateg. Inf. Syst.* **2017**, *28*, 66–82. [CrossRef]
42. Forrester, J. *Industrial Dynamics*; Massachusetts Institute of Technology Press: Cambridge, MA, USA, 1961.
43. Sterman, J.D.; Meadows, D. STRATAGEM-2. *Simul. Games* **1985**, *16*, 174–202. [CrossRef]
44. Duggan, J. An Introduction to System dynamics. In *System Dynamics Modeling with R*; Springer: Cham, Switzerland, 2016; pp. 1–24. [CrossRef]
45. Pruyt, E. *Small System Dynamics Models for Big Issues: Triple Jump towards Real World Complexity*; TU Delft Library: Delft, The Netherlands, 2013; 324p.
46. Lane, D.C. On a resurgence of management simulations and games. *J. Oper. Res. Soc.* **1995**, *46*, 604–625. [CrossRef]
47. Meadows, D. A brief and incomplete history of operational gaming in system dynamics. *Syst. Dyn. Rev.* **2007**, *23*, 199–203. [CrossRef]
48. Cunico, G.; Aivazidou, E.; Mollona, E. System dynamics gamification: A proposal for shared principles. *Syst. Res. Behav. Sci.* **2021**. preprint. [CrossRef]
49. Papathanasiou, J.S.; Armenia, S.; Barnabè, F.; Carlini, C.; Ciobanu, N.; Digkoglou, P.; Jarzabek, L.; Kulakowska, M.; Lanzuisi, A.; Morfoulaki, M.; et al. Game Based Learning on Urban Sustainability: The "Sustain" Project. In Proceedings of the 11th International Conference on Education and New Learning Technologies, Palma, Spain, 1–3 July 2019. [CrossRef]
50. Connolly, T.M.; Boyle, E.A.; MacArthur, E.; Hainey, T.; Boyle, J.M. A systematic literature review of empirical evidence on computer games and serious games. *Comput. Educ.* **2012**, *59*, 661–686. [CrossRef]
51. Qudrat-Ullah, H. Perceptions of the effectiveness of system dynamics-based interactive learning environments: An empirical study. *Comput. Educ.* **2010**, *55*, 1277–1286. [CrossRef]
52. Martinez-Moyano, I.J.; Conrad, S.H.; Andersen, D.F. Modeling behavioral considerations related to information security. *Comput. Secur.* **2011**, *30*, 397–409. [CrossRef]
53. Bier, A.; Anderson, B. Cooperation and Learning in Cyber-security Training Exercises. In Proceedings of the 31st International Conference of the System Dynamics Society, Cambridge, MA, USA, 21–25 July 2013.
54. Armenia, S.; Franco, E.F.; Nonino, F.; Spagnoli, E.; Medaglia, C.M. Towards the Definition of a Dynamic and Systemic Assessment for Cybersecurity Risks. *Syst. Res. Behav. Sci.* **2018**, *36*, 404–423. [CrossRef]
55. Zeijlemaker, S. Unravelling the Dynamic Complexity of Cyber-Security: Towards Identifying Core Systemic Structures Driving Cyber-Security Investment Decision-Making. Ph.D. Thesis, Radboud University, Nijmegen, The Netherlands, 16 March 2022.
56. Andersen, D.; Moore, A.P.; Stanton, J.M.; Cappelli, D.M.; Rich, E.; Weaver, E.A.; Gonzalez, J.J.; Sarriegui, J.M.; Zagonel, A.; Mojtahedzadeh, M.; et al. Preliminary System Dynamics Maps of the Insider Cyber-threat Problem. In Proceedings of the 22nd International Conference of the Systems Dynamics Society, Oxford, UK, 25–29 July 2004; pp. 1–36.
57. Armenia, S.; Cardazzone, A.; Carlini, C. Understanding Security Policies in the Cyber Warfare Domain through System Dynamics. In Proceedings of the 4th International Defense and Homeland Security Simulation Workshop (DHSS 2014), International Multidisciplinary Modeling and Simulation Multi-conference (I3M 2014), Bordeaux, France, 10–12 September 2014.
58. ISACA. *CISM Review Manual 2015*; ISACA: Rolling Meadows, IL, USA, 2015.
59. Baskerville, R.; Spagnoletti, P.; Kim, J. Incident-centered information security: Managing a strategic balance between prevention and response. *Inf. Manag.* **2014**, *51*, 138–151. [CrossRef]
60. Böhme, R.; Moore, T. The Iterated Weakest Link, a Model of Adaptive Security Investment. *J. Inf. Sci.* **2016**, *7*, 81–102. [CrossRef]
61. Su, X. *An Overview of Economic Approaches to Information Security Management*; University of Twente, Information System Group: Enschede, The Netherlands, 2006.
62. Chismon, D.; Ruks, M. *Threat Intelligence: Collecting, Analysing, Evaluating*; MWR Info Security: Basingstoke, UK, 2015.
63. Syed, R. Cybersecurity vulnerability management: A conceptual ontology and cyber intelligence alert system. *Inf. Manag.* **2020**, *57*, 103334. [CrossRef]
64. Vogus, J.T.; Sutcliffe, K.M. Organisational resilience: Towards a theory and research agenda. In Proceedings of the IEEE International Conference on Systems, Man and Cybernetics, Montréal, QC, Canada, 7–10 October 2007.
65. Linkov, I.; Bridges, T.S.; Creutzig, F.; Decker, J.; Fox-Lent, C.; Kröger, W.; Lambert, J.H.; Levermann, A.; Montreuil, B.; Nathwani, J.; et al. Changing the resilience paradigm. *Nat. Clim. Chang.* **2014**, *4*, 407–409. [CrossRef]
66. Reinmoeller, P.; Baardwijk, N. The Link between Diversity and Resilience. *MIT Sloan Manag. Rev.* **2005**, *46*, 60–65.
67. Scott, R.J.; Cavana, R.Y.; Cameron, D. Recent evidence on the effectiveness of group model building. *Eur. J. Oper. Res.* **2016**, *249*, 908–918. [CrossRef]
68. Ford, D.N.; Sterman, J.D. Expert knowledge elicitation to improve formal and mental models. *Syst. Dyn. Rev.* **1998**, *14*, 309–340. [CrossRef]
69. Forrester, J.W.; Senge, P.M. Tests for building confidence in system dynamics models. *TIMS Stud. Manag. Sci.* **1980**, *14*, 209–228.
70. Barlas, Y. Formal Aspects of Model validity and validation in system dynamics. *Syst. Dyn. Rev.* **1996**, *12*, 183–210. [CrossRef]

71. Liang, T.Y.; Ta, C.K. Strategic information technology plan: A vital component in the corporate strategies of banks. *Inf. Manag.* **1994**, *26*, 265–272. [CrossRef]
72. Montazemi, A.R.; Qahri-Saremi, H. Factors affecting adoption of online banking: A meta-analytic structural equation modeling study. *Inf. Manag.* **2015**, *52*, 210–226. [CrossRef]
73. *European Network and Information Security Directive 2016/1148 (NIS 2.0)*; European Commission: Luxembourg, 2016.
74. *Digital Operational Rescilience for Financial Services 2020/0266 (DORA)*; European Commission: Brussels, Belgium, 2020.
75. Desouza, C.K. Strategic contributions of game rooms to knowledge management: Some prelimenary insights. *Inf. Manag.* **2003**, *41*, 63–74. [CrossRef]
76. Woolley, A.W.; Chabris, C.F.; Pentland, A.; Hashmi, N.; Malone, T.W. Evidence for a Collective Intelligence Factor in the Performance of Human Groups. *Science* **2010**, *330*, 686–688. [CrossRef]
77. Tversky, A.; Kahneman, D. Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychol. Rev.* **1983**, *90*, 293–315. [CrossRef]
78. Park, C.; Whan, L.; Parker, V. Familiarity and Its Impact on Consumer Decision Biases and Heuristics. *J. Consum. Res.* **1981**, *8*, 223–230. [CrossRef]
79. Slovic, P.; Finucane, M.L.; Peters, E.; MacGregor, D.G. The affect heuristic. *Eur. J. Oper. Res.* **2007**, *177*, 1333–1352. [CrossRef]
80. Rao, H.; Greve, H.R.; Davis, G.F. Fool's Gold: Social Proof in the Initiation and Abandonment of Coverage by Wall Street Analysts. *Adm. Sci. Q.* **2001**, *46*, 502–526. [CrossRef]
81. Festinger, L. *A Theory of Cognitive Dissonance*; Row & Peterson: Evanston, IL, USA, 1957.
82. Seligman, M.P.; Maier, S.F. Failure to escape traumatic shock. *J. Exp. Psychol.* **1967**, *74*, 1–9. [CrossRef]
83. Grossklags, J.; Christin, N.; Chuang, J. Predicted and observed user behaviour in the weakest-link security game. In Proceedings of the Workshop on Usability, Psychology, and Security, San Francisco, CA, USA, 14 April 2008.
84. Grossklags, J.; Reitter, R. How Task Familiarity and Cognitive Predispositions Impact Behaviour in a Security Game of Timing. In Proceedings of the IEEE 27th Computer Security Foundations Symposium, Vienna, Austria, 19–22 July 2014.
85. Nochenson, A.; Grossklags, J. A behavioural Investigation of the Flipit Game. In Proceedings of the Workshop on the Economics of Information Security, Washington, DC, USA, 11–12 June 2013.