Introducing Resource Management Into a Large-Scale Cyber Security Seminar Game: A Case Study

Peadar Callaghan
Tallinn University, Estonia

peadar@tlu.ee

Abstract: In seminar wargames, participants are asked to resolve problems presented by an umpire, who also adjudicates the outcome of the player decisions. This type of game has been used across a wide range of contexts, including notably in cybersecurity training. Seminar games, however, have been criticized for their lack of rigor. The solutions developed during the game can often be idealized. It is also common for these games to have pre-scripted outcomes. This minimizes the agency of the player and reduces the investment the player can have in the game. This paper focuses on the design and development of a seminar wargame with a novel resource management component. This component expands on the traditional seminar game and addresses some underrepresented training needs. The game was designed using a participatory design framework and the process is clearly described in the paper. The game was then played during a large-scale cybersecurity event by governmental representatives from 18 different countries using a hybrid game system. Data was gathered through a series of observations made during play, interviews with participants, and discussion during a debriefing session. Preliminary results show that the participants and sponsors felt that the game improved the overall quality and utility of the seminar style wargame. While the practical implementation of the game requires refinement, the overall approach is viable for a broad range of contexts in which seminar-style games can be used.

Keywords: Cybersecurity, Training, Participatory design, Game-based learning, Seminar game, Wargame

1. Introduction

The field of cybersecurity is one of continued growth. There is an estimated shortfall of over 3 million jobs worldwide in the sector for 2023 (Blažič, 2021). Attempts to address this shortfall have ranged from the introduction of degree-level courses and summer camps to the use of games in teaching and training (Jin, *et al.* 2018). These initiatives have tended to focus on the technical training of future cybersecurity professionals, utilizing approaches such as capture the flag competitions and cyber ranges to allow technical students to practice in a realistic manner. However, cybersecurity exists across several domains, not just the technical. From the lowest-level employees to the high-level government decision-makers, an understanding of cybersecurity is becoming essential. Unfortunately, the training used for these groups often focuses on old school chalk and talk modalities that can do little to actually change their behaviour.

Traditionally, the games used to train people involved in the strategic decision-making domain have been a form of seminar wargame game. Originally developed by the Rand Corporation in the 1950s, seminar games sit the players around a table and present them with dilemmas. Solutions are described by the participants and the umpire judges their effectiveness. Then the participants are given another dilemma (Perla, 1985). While these games have demonstrated their utility in a range of contexts from predicting future oil values (Bradfield et al., 2005) to the impact of nuclear war (Kahn, 1984), they have also been criticized for their lack of rigor. It is common for these games to have pre-scripted outcomes, reducing player engagement due to the limited agency and potentially reinforcing pre-existing prejudices instead of providing novel solutions. Further, the solutions developed during the game can often be idealized and lack consideration for the practical constraints of real-world cyber warfare, such as limited resources. This has led some experts to derogatorily label these games as BOGSAT's: a Bunch of Guys Sat Around a Table (Caffrey, 2019).

To address this problem, a new cybersecurity training game was designed. This game introduced a resource management component to a hybrid analogue-digital cybersecurity seminar-style wargame aimed at governmental representatives taking part in a large-scale international cybersecurity exercise. By adding resource management, the game introduces the importance of task prioritization when operating with a limited pool of resources amid a simulated series of crises. This was done to increase the realism, as well as players' investment in and engagement with the game.

2. Methodology

2.1 Context of use

The game was designed to be a part of a large-scale (over 1,500 participants) cybersecurity exercise known as Locked Shields. Other parts of the overall exercise would include a technical game using a cyber range. Two other seminar games were also run in parallel, one focusing on the legal aspects of cyber incidents under state

and international law, and the other examining disinformation and how to construct media policy. With the game being part of such a large-scale event, it was essential to get all aspects of the game approved by multiple stakeholders throughout the process.

2.2 Research Design

This study used a participatory design methodology: an inclusive design methodology that seeks to draw experiences from the end user of a product in its creation (Simonsen and Robertson, 2012). A key element of participatory design is that it focuses on the production of an end product, system or artifact. This was especially suited to the purpose of this study as a game had to be produced by the deadline. The other advantage of participatory design is that it encourages buy-in from multiple stakeholders. This makes the approval process for the adoption of the final product more straightforward.

The participants involved in the design process included several experts in the field of cybersecurity policy and strategic decision making from a recognized centre of excellence dedicated to the subject; a player of the seminar game from the previous year's game used at the same event; and an expert in the design of the computer platform used to facilitate the game. The author of the paper acted as a facilitator and game design consultant.

The first step was to conduct an ideation and brainstorming session with key participants. Following this, rapid iterative prototyping was used. This enabled all participants to provide feedback and make suggestions for improvement throughout the process. The game went through eight separate iterations of design, involving five training sessions and three full playtests with players who had no experience with the design, before it was deployed as a final product.

Qualitative data was gathered during both the playtesting and the final deployment of the game based on a combination of observation and interview/discussion with the players. Observations were conducted after every resource management turn during playtesting and after every 10 minutes during the implementation of the game. Observations focused on the game state of the various participants. Notes were made on decisions that were being observed and perceived difficulties being experienced.

Informal interviews were conducted with the playtesting groups to gather ideas for improvements to be made. A semi-structured interview based on the recommended debriefing method for wargames (Van den Hoogen, Lo and Meijer, 2016) was used with players of the final game. Finally, a debrief was held on the day after the final play concluded. Impressions and comments were sought from the participants through an open discussion.

3. Design and Implementation

3.1 Game Concept Outline

Before commencing the design, the seminar game from the previous year was analysed to generate a clear view of the limitations to the game design and understand the audience and context of use for the game. Below is an outline of its gameplay, which became a starting point for the redesigned game.

The game would take 3 hours to play and involve two types of players. These were the white team liaisons (WTLs) and the strategic teams (ST). WTL's were representatives of the country or team who were engaging with the training game. They would travel to the physical location of the exercise and coordinate the play with the strategic teams located back in their home country.

WTL's are trusted agents within the system. This means that they have more information than other players and a responsibility to confidentiality. This is similar to how a dungeon master may work in roleplaying games. They have access to all the information of the setting and adventure (scenario and inject decks containing sequenced in-game events) so that they can advise the strategic team on what roles will be needed. At the same time, they have a responsibility to not spoil the game by giving too much away. Their position as mediators helps to keep the strategic teams on track and provides translation and other skills that allow the game to be internationally playable. The WTL's were involved in all aspects of the game design and testing.

The strategic teams (ST's) were made up of representatives of national governments. These people would be expected to either make decisions or advise political leaders during a cyber security crisis. They could be drawn from the civil service, military, or critical industries. The makeup of the teams would not be known by the designers before gameplay and the identity of the players would be kept confidential throughout the process.

During the game, the WTL would ask the players to imagine their country was undergoing a series of cyber-attacks, narrate the events and posing questions for the ST to discuss and answer. Given that questions used in the game could touch on areas of governments security preparedness for crisis, the answers were highly sensitive. Due to the ST's being located in their home countries, their interactions with the WTL's would take place online using a dedicated software system. This would mean that the game system would have to be simple enough to be used with minimum support from the designers.

3.2 Ideation

The initial brainstorming and ideation meeting took place in late October 2022. The discussion highlighted the disappointment of stakeholders with the previous years' game. It was decided that the new design should act as an add-on to a traditional seminar game rather than replacing it, increasing the value of playing the game and making it more engaging.

Three key issues were identified during the discussion: 1) the lack of player agency and game persistence (each of the decision points was independent from all other decisions), 2) an assumption that the solutions presented by players were to take effect immediately, and 3) strategic decisions being taken in isolation from their operational context. The game design would thus focus on addressing these three core issues.

3.3 Prototyping

Once these key issues and practical constraints were defined, a prototyping session was held in November at Tallinn University's Game Lab. During this session, rapid prototyping was used to address the issues identified in the ideation meeting.

To address the lack of agency, players were given a series of projects that they could choose to start or pause at any time during the game. These projects would be linked to the strategic questions that they had already engaged with in the regular seminar side of the game. To deal with the immediacy issue, a turn-tracking mechanic was introduced, whereby the projects would take several turns to complete. Finally, there needed to be resources that could be used to begin projects. These resources would have to be limited to make the decision of which project to begin meaningful for the player and reflect the limited number of skilled cyber-operatives any government may call on. These three components and their interactions became the core of the gameplay.

The updated game involved two distinct types of questions. Type one were the traditional strategic-level questions typical of seminar games. These would trigger discussion and an answer, but would not trigger the decision to start a project or not. The other type of question was "actionable," meaning it would trigger the creation of a project. If the players ran out of resources but wanted to start a new project, they would have to either pause a current project or wait for a project to reach completion. The core loop of the game is summarized in Figure 1.

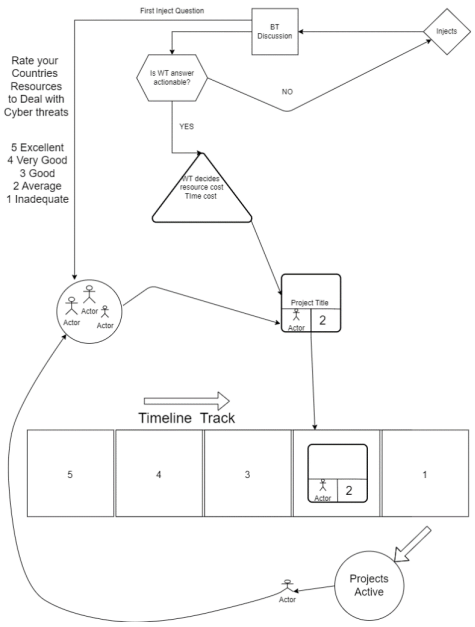

Figure 1: Core gameplay loop

Below is an outline of other major changes introduced into the game across several iterations of prototyping.

Hybrid board game design

Originally, the game was to be computer-based so that everyone could track the progress of projects and see the state of play. After consulting with the commercial supplier of the software used in the event, it was decided that the timeline was too short to implement such as system. However, they would work on rolling it out for the 2024 edition. In the meantime, the 2023 version would be designed as a board game controlled and administered by the WTLs, who would communicate the game information to the ST participants online.

Resources and turn-tracker

Originally, resources were conceived as being between 2 and 10. They would represent the teams that the players had available to address the problems. The decision to deploy the resources or not would be the key decision point in the resource management game. While governments have access to many resources, it was decided to limit them for game purposes to help to create dilemmas for the players to explore. It was also decided early in the design process not to have teams with different specializations, as this would further complicate the game and add to the processing load on the players for a limited educational return.

The allocation of resources to the teams went through several iterations and discussions. Initially, the plan was to ask each strategic team to rate their countries readiness on a scale from 1 (poor) to 5 (excellent); then, two teams would be awarded for each point the country had given. This was changed in part because some countries have historically overemphasized their capabilities, while others may underplay their capabilities in a cultural form of modesty. The next iteration focused on asking players questions drawn from the Cyber Readiness Index (Gomez, 2013), which would focus on specific competencies and result in additional teams being awarded in addition the baseline of three teams. Eventually, it was decided to simplify it still further and give only four teams to each ST. This was done so that the exhaustion of resources would occur faster in the game, forcing decisions to be made and ensuring that not every project could be started.

This decision also affected the turn-tracker on the game board. As the game was originally conceived as being computer-mediated, the timeline for each project was not originally fixed. However, due to the above considerations and practical constraints such as the small size of the game board (A4), the number of turns on the turn tracker timeline was fixed at four.

Project cards

Two different types of project cards were defined: Standard and Mandatory projects. In Mandatory projects, the player had to begin the project no matter what. The initial Mandatory project was conceived as a flu outbreak within the cyber-response teams that may have required staff reallocation for a set number of turns. Standard projects were to be given a cost in number of teams by the WTL and they would also decide how long they felt the project would take to reach completion. During the online training sessions, this was found to be too confusing for the WTL's and led to the collapse of one of the test runs. While the initial system would have provided more flexibility and provided a more tailored experience, it came at the cost of complexity in training and execution.

To resolve this issue, project cards were simplified to being only one team per project and the time to completion being fixed at four turns for both Standard and Mandatory projects. This had a knock-on effect on the resources, as the number of teams would have to be limited to no more than four to prevent the strategic teams from completing every project they were presented with.

Inject cards

In order to compensate for the insufficient amount content in previous years' games (players would often exhaust the content before the three-hour playtime ended), 70 inject cards representing in-game events were designed. During playtesting, it was discovered that this was too much material, and it had to be dramatically cut down for the final game, resulting in 36 inject cards. The cards were printed with key text about the inject and the number in which the card was to be played, as the in-game events were designed to take place sequentially.

Miscellaneous changes

The original game board used for the full playtest is presented in Figure 2. The aesthetics of the game board were subsequently changed to conform with the corporate identity of the larger exercise. The colour and

markings of the board were also changed, and the logos of the exercise were added (the final version will not be shared in this paper).

The original tokens representing the teams were redesigned to stand more upright and be easier to handle. The tokens were designed in Blender and then 3D printed.

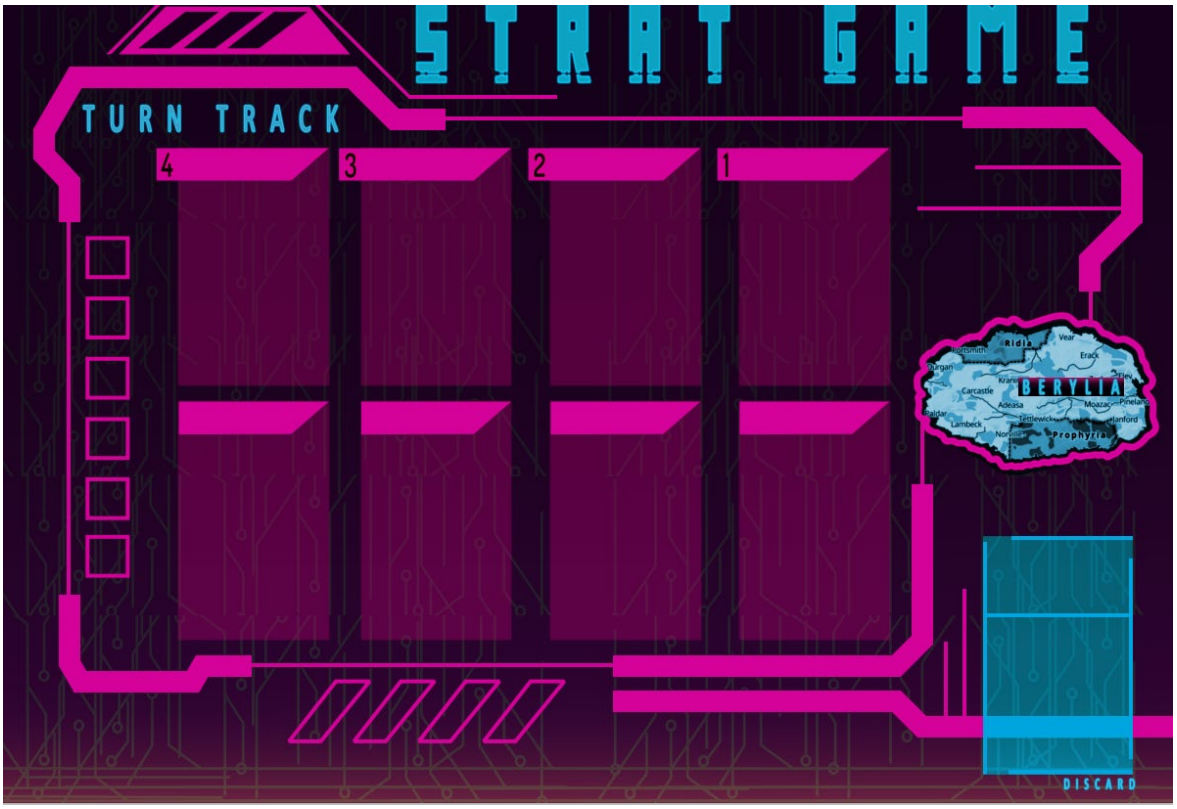

Figure 2: The original game board

3.4 Final Ruleset

The following rules were used in the final version of the game.

Setup phase

- Place one team token (four in total) in the squares on the left-hand side of the board.
- Place the project deck face down to the left of the tokens.
- Reveal the first card.
- If the players decide to send a team to support their ally, place one team token in the country on the far-right hand side of the board.
- Otherwise, place the card on the bottom left-hand square of the game board.

Decision phase

- During the resource management phase, the next project card is revealed.
- If the players wish to begin the project, they must staff it with a team.
- If there are not enough teams available, a team may be pulled from another project.
- A player may decide to staff any project card that is currently available.
- Mandatory projects must be staffed and completed.

Movement phase

- Once all decisions have been made, the movement phase begins.
- Any unstaffed projects cards are moved down to the row beneath their current position.
- Any staffed projects move one step over to the right.
- Any new or restarted projects move up to the active row.

- If a project reaches the end of the active row, the project is complete, and the team returns to the available position.

3.5 Running the Game

The strategic seminar game included representatives from 18 separate countries spread across four continents and one team of industry specialists. Each team had a white team liaison in the game room located in Tallinn in Estonia. The WTL would act as mediator and controller for the pace of the exercise. In addition, they had an open channel of communication to their home countries' team to facilitate the gameplay process. Answers to all the questions were gathered on a dedicated secure server maintained by a company specializing in cybersecurity exercises.

The game was played for 3 hours with full technical support on standby. Given that players were engaged in multiple time zones, the game network was left open for two days to allow teams to continue play. The extra playtime did not include rapid technical support or observation of play, however.

4. Results and Discussion

4.1 Observation

18 out of the 19 WTLs engaged with the resource management game. The one who chose not to engage had not attended the trainings for the system online and even with support given the day before was not comfortable with the system.

An additional team decided to play the game out of sequence and readjusted their deck to suit this new ordering. The game was robust enough to handle these changes, although this did throw off the intended use of the mandatory projects to force the players to have all teams active. The WTL did not report any problems experienced because of this.

The very first decision presented to the players was whether to send a team to aid an ally who was experiencing an attack. All ST's sent a team to aid in the turn: 17 did so when it was initially requested, and the final ST sent the team after almost two hours of play, when several projects at home had been completed. At no point did any players pull their team from supporting their allies. This could be showing the strength of commitment of the nations involved to mutual defense, or it could show that they were not put under sufficient pressure by the events being simulated in their home country. Either way, it is an interesting data point to explore in future iterations.

One of the mandatory events was a call from the prime minister's office asking for a team to be devoted to look into an embarrassing hack. One team decided to ignore this completely and refused to begin the project for one turn until their WTL insisted. The liaison theorized that this was because his strategic team was made up of primarily technical people, who did not see the necessity of the project.

One WTL brought a specific webcam and tripod setup so that they could display the resource management board to the strategic team at all times. This was in line with what had been done during the training sessions for the WTL's. He reported that it made explaining the situation a lot clearer, marking this as a useful approach for future implementation.

Two distinct play styles were observed when dealing with the resource management component. The first could be labelled the conservative style. When playing conservatively, players would always have a team in reserve ready to be deployed should an emergency arise. They would often have a backlog of projects and would decide on the project to begin based on the importance that they assigned them. It was rare for them to begin a project when it was first presented to them; instead, they would wait to see what was happening next before choosing to act. The second, more aggressive style of play, was to have teams working on projects at all times. This form of play emphasized the ability to pull teams out of projects in case of emergencies or mandatory projects. Trace data of actual game play will have to be reviewed to see if one strategy produced better results than the other. In general, the team drawn from the industry played much more aggressively, while those teams who reported having military involvement played much more conservatively. A similar pattern had been observed in playtesting, where the student players (the non-experts) were more aggressive in their approach than the expert team. How much this reflects cultural practices inside the players is uncertain, but would be interesting to investigate further.

One strategic team continued to play even after the time had expired and completed all injects with a full discussion of the relevant information. This took over 8 hours of full play time, showing significant engagement and commitment.

4.2 Interviews and Debrief

Four interviews were conducted with WTL's immediately after the game. All participants expressed a positive view on the value of the game. One interviewee reported that their strategic team spent almost equal time debating the use of resources as they did the strategic questions. This was a surprising finding, as the resource management component is a simple system of binary choices (staff the project or don't staff the project). This highlights that the introduction of resource management increased both the player engagement and raised interesting problems for the player.

The debrief was held the day after the game. Its key points are discussed below.

The primary emotion experienced during the game for the WTL's was one of stress. They felt they had a lot of things to be aware of and to be responsible for. With deeper probing, however, this was found not to be negative, as most WTL reported also finding the experience interesting and valuable. Several WTL's reported that the time seemed to go by quickly, which is suggestive of entering a flow state (Chen, 2007). Further, this somewhat contradictory reporting is consistent with the concept of games' capacity to elicit a state of pleasurable frustration (Gee, 2013).

Even with the reported stress, there was a great level of enthusiasm for the game. Several liaisons suggested expanding the play to take place over two days rather than just one. The suggestion was also to split the players into senior decision-makers and lower-level, on-the-ground decision makers on the second day. The splitting of the strategic teams was hotly debated, and a consensus was not reached in the room.

There was a call for the board to be virtualized. It was felt that this would help with the play, as then everyone could see the game state clearly. In addition, it would allow visualizations of progress or lack of progress to be used in analysis. The virtualization would also help to address the next issue raised.

The participants reported a lack of a sense of consequences for their decisions. If they completed a specific project, they were not sure of the implications on the grander scale. This betrayed the lack of a fail state in the game as it was originally designed and a lack of established ludic goal. In the absence of these, teams were free to create their own interpretations of the goals, but this was at the cost of a sense of the decisions being important to the gameplay. In the future, negative gameplay feedback may be introduced to see how the situation deteriorates if crucial decisions are not made or followed through on.

Some concerns with the realism of the projects were expressed, in particular concerning the banking sector, which some of the projects related to. In reality, the banking sector has more cybersecurity resources at its disposal than many national governments. As such, it may have been unrealistic to create any projects where a national team would step in to provide cybersecurity support.

One of the major recurrent criticisms of seminar wargames—their often linear progression—emerged in the discussion. There were calls to create multiple pathways through the inject deck, leading to a branching narrative. This would increase engagement and allow strategic teams with different strengths to work on problems in parallel. The branching structure would allow the strategic teams to prioritize the questions they felt were important and to have more agency over the game experience. This is an area of potential future research and design.

5. Conclusion

To the author's knowledge, this is the first example of resource management being used in a non-classified seminar wargame. It serves as a proof of concept for the inclusion of these elements to increase both engagement and training value. The various stakeholders have expressed an interest in including this method in further trainings. This has included representatives from industrial partners, who wish to develop the method into a product, as well as members of various military training institutions and centers of excellence. As such, the utility and practicality of the approach has been recognized by both players and game sponsors.

Overall, the design process has been one of initial complexity moving towards simplification. The instinct of the subject matter experts was to attempt to model reality as closely as possible. Yet, the objective of the game was not to accurately reflect the world, but to raise interesting questions for discussion, while remaining understandable to its players.

While game trace data was recorded during the full game session for all 19 teams, the data requires a great deal of processing before it can be examined and publicized, given the sensitive nature of the subject matter. When analyzed, the trace data is expected to provide information on strategic prioritization, demonstrating which projects the ST's felt they had most responsibility for and which they saw as being the responsibility of other sectors. The data could be of use to discussions of policy and strategic studies, as well as to inform future game designs.

The mechanics of this game can be used in other contexts to examine a wide range of topics. Another game using a similar system is currently scheduled for June 2023 and will examine trust in cyber-crises. Futures games can use resources to measure different metrics from money to teams to bandwidth allocation, all of which increase the value of the seminar game.

References

Blažič, B.J. (2021) 'The cybersecurity labour shortage in Europe: Moving to a new concept for education and training', *Technology in Society*, 67, p. 101769. Available at: https://doi.org/10.1016/j.techsoc.2021.101769.

Bradfield, R. *et al.* (2005) 'The origins and evolution of scenario techniques in long range business planning', *Futures*, 37(8), pp. 795–812. Available at: https://doi.org/10.1016/j.futures.2005.01.003.

Caffrey, M.B. (2019) *On wargaming: how wargames have shaped history and how they may shape the future*. Newport, Rhode Island: Naval War College Press (Naval War College Newport papers, 43).

Chen, J. (2007). Flow in games (and everything else). *Communications of the ACM*, 50(4), pp.31-34. Available at: https://doi.org/10.1145/1232743.1232769.

Gee, J.P. (2013) *Good video games and good learning: Collected essays on video games, learning and literac*y. New York: Peter Lang.

Gomez, A. (2013) 'Cyber Readiness Index 1.0', *Science, Technology, and Public Policy Program, Belfer Center*, p. 8.

Jin, G. *et al.* (2018) 'Evaluation of Game-Based Learning in Cybersecurity Education for High School Students', *Journal of Education and Learning (EduLearn)*, 12(1), pp. 150–158. Available at: https://doi.org/10.11591/edulearn.v12i1.7736.

Kahn, H. (1984) *Thinking about the unthinkable in the 1980s*. New York: Simon and Schuster.

Perla, P.P. (1985). What Wargaming is and is Not. *Naval War College Review*, 38(5), pp.70–78. Available at: https://www.jstor.org/stable/44637075.

Simonsen, J. and Robertson, T. (eds) (2012) *Routledge International Handbook of Participatory Design*. New York: Routledge. Available at: https://doi.org/10.4324/9780203108543.

Van den Hoogen, J., Lo, J. and Meijer, S. (2016) 'Debriefing Research Games: Context, Substance and Method', *Simulation & Gaming*, 47(3), pp. 368–388. Available at: https://doi.org/10.1177/1046878116651023.