Siguru: A Case Study

Gamification of Cybersecurity Awareness Training Materials

Ioan Thomas, BSc

Original thesis supervised by
Marcus Nohlberg, PhD

January 19, 2025

Austria

research@ioan.at

Contents

Chapter 1

Introduction

In recent years, gamified[1] awareness training materials have gained significant attention, not only in educational settings but also in corporate environments (e.g. reflected in articles (Anadea, 2018) and products/services (PricewaterhouseCoopers, n.d.)) where employees must be educated on complex topics such as information security (van den Boer, 2013). These games aim to create a dynamic and interactive experience, in which participants engage more easily and actively than with more passive traditional methods. Backing the rising popularity of well-designed gamified learning in e-learning and educational settings, studies have demonstrated positive effects on participants (e.g. Jia et al. (2016)). Attributes which may apply to gamified awareness training materials as well.

However, despite the increasing adoption of gamified formats, these methods remain relatively underexplored from a scientific perspective. Rigorous research is required to validate experimental tools to ensure they are effective. Particularly for fields such as cybersecurity, where effective training and a secure intuition may reduce risk. Previous research focuses on educational settings such as schools and universities, with fewer studies addressing awareness training on specialised topics.

This paper, originally part of a full thesis tackling gamification of security awareness materials (Thomas, 2024), covers a standalone case study of Siguru, a game created by Nohlberg (2024). One of the original goals of the full thesis was to measure player enjoyment, engagement and the overall learning experience by creating or adapting a system to do so. The goal of this adapted case study is to assess the game's capacity to engage users, communicate its core messages, and promote behavioral change or enhanced understanding in the target audience. The evaluation is based mainly on quantitative data gathered from a questionnaire but also considers qualitative insights derived from player feedback during playtests.

[1] Gamification is the use of game elements and game thinking in non-game environments to increase target behaviour and engagement. (van den Boer, 2013)

The primary research questions guiding this study are:

1. How effective is the game in fostering engagement amongst players?
2. What is the player's perception after participating in this game?
3. How well does an existing tool for evaluating e-learning materials perform when applied to Siguru?

This study aims to provide an analysis of Siguru by Nohlberg (2024) and an analytical tool to evaluate awareness training games.

Chapter 2

Case Study Siguru

Siguru is a cybersecurity card game created by Marcus Nohlberg, a professor with a research focus on the human aspects of information security, with the goal of teaching players information security concepts and terminology by versing teams of two against each other. Players need to understand and explain information security attacks and defenses to win or use the provided game mechanics to gain an advantage that might compensate for gaps in knowledge. The game provides a deck of cards containing terminology cards, similar to flash cards, with varying levels of difficulty and obscurity. Its website offers a point counting system and a timer to keep track of game progression. The cards also display a QR code which leads players to explanations of the rules, action cards and gives players the opportunity to educate themselves and look up answers to cards and terminology they may not be familiar with. (Nohlberg, 2024)

The initial play test is conducted and evaluated in the following sections with a brief explanation of the contents and mechanics of the game.

2.1 Play Test Procedure

For the pilot play test the group of students with a cybersecurity focus were presented the game, the rules, the point system and the instructions on how to play by Marcus Nohlberg. The players already having registered as teams or showing up spontaneously and forming a new team received a deck of cards and familiarised themselves with the game. Marcus Nohlberg and I separated our responsibilities by having Marcus respond to any questions about the game and its rules while all feedback regarding the game should be directed at me for documentation. This way we ensured that a degree of scientific rigour would be upheld. Additionally, any passive observations, e.g. complaints, confusion, frustrations, compliments, were observed and documented. After completing the pilot tournament in which 3 matches were played per team and their points added to a grand total with a winning team declared the participants were asked to complete an evaluation questionnaire.

2.2 Evaluation method

To evaluate the game a questionnaire was chosen as an accessible and easy method for the participating students. The goal is to understand how well the gamified format performed with the players and the overall enjoyment of the game itself. To be able to measure this a questionnaire created by Fu et al. (2009) with slight adaptions was chosen which are covered in Subsection 2.2.1. Additionally, observations during the active tournament were noted to provide some qualitative data as an additional insight to ensure that feedback not shared in the questionnaire is documented too.

2.2.1 Questionnaire Basis

As effective gamified learning materials share many common requirements with effective e-learning games, with the main difference being the digital medium explicitly specified by mentioning "e-learning", the EGameFlow scale created by Fu et al. (2009) could be used as a tool to measure learners' enjoyment of educational materials in a gamified format. EGameFlow covers a variety of topics related to gamified learning materials with statements to be evaluated on a Likert-Scale from 1 (lowest) to 7 (highest). Its intention is enabling the developer of a game to view their content from the learner's and player's perspective. Being aware of strengths and flaws of learning games can help improve and boost the fulfillment of learning objectives. The authors of this EGameFlow scale have verified its reliability and validity by distributing the questionnaire/scale to various students spread across different games enrolled in a technical learning course resulting in 166 valid samples. (Fu et al., 2009) The scale consists of eight dimensions which can be identified by the letters in parentheses:

Immersion *(I)*: The game should immerse players and keep them engaged

Social Interaction *(S)*: Tasks should form an environment for social interaction between players

Challenge *(H)*: The game should offer challenges to different skill levels that may be adjusted accordingly

Goal Clarity *(G)*: Tasks and rules need to be clearly explained in the beginning

Feedback *(F)*: Players can assess their current knowledge status and compare it to an ultimate understanding of the entire content

Concentration *(C)*: Game activities should enable concentration

Control/Autonomy *(A)*: Players should be in control of their actions and choices and enable taking the initiative

Knowledge Improvement *(K)*: Knowledge and skill of the players should increase by playing the gamified learning materials

(Fu et al., 2009)

Many of these dimensions can be mapped to background research conducted in the original thesis. For example Furdu et al. (2017) explains the importance of feedback and Chou (2015) covered the wide range of game factors, of which many can be found under the same or similar name. (e.g., An, 2020; Chou, 2015; Furdu et al., 2017)

2.2.2 Adapted Questionnaire

During the evaluation process of reliability and validity for the EGameFlow scale by Fu et al. (2009) some items were marked for deletion. These items are excluded from the questionnaire used in the Siguru game evaluation and the final visual analogue "overall sense of enjoyment" scale is simplified to a common star rating system from 0 to 5 describing the overall rating. Additionally, the questionnaire based on the EGameFlow scale is expanded by 16 additional items directly relating to game drivers suggested by Chou (2015) and the summarising analysis in a literature overview by An (2020) to gain a deeper understanding of the game design and additional feedback from the player's point of view. These additional entries were categorised as *Game Drivers* with the category identifier being *(GF)*. Finally, players are able to fill out an open-ended question regarding any additional feedback they may want to share.

The full questionnaire can be referenced in the Appendix A.3.

Chapter 3

Evaluation of Siguru

After the initial play test the participants, consisting mainly of information security students, were asked to fill out the questionnaire. From the 15, later 13, players 8 valid samples were collected and evaluated.

3.1 Results of the questionnaire

3.1.1 Parsed Data

ID	Mean	Median	Standard Deviation	Range
C1	5	5	1.4142	3-7
C2	5.125	5	1.2464	3-7
C3	5.375	6	1.8468	2-7
C4	4.875	5	1.6421	2-7
C5	5.375	6	1.685	2-7
C6	5.75	6	1.2817	4-7
G1	5.25	6	2.1213	2-7
G2	4.25	4.5	1.8323	2-7
G3	4.625	5	1.1877	3-6
G4	4.375	4	1.5059	3-7
F1	4.625	5.5	2.3261	1-7
F2	4.375	4	1.1877	3-7
F3	3.75	4	2.0529	1-7
F4	3.875	3.5	2.031	1-7
H1	3.375	3.5	1.4079	1-5
H2	4.875	5.5	1.9594	2-7
H3	2.5	2.5	1.6903	1-6
H4	3.25	3.5	1.9821	1-7
H5	5	5.5	1.5119	3-7
H6	5.5	6	1.8516	2-7
A1	5.5	6	2	1-7
A2	5.125	6	1.8851	2-7

Table 3.1 (*continued*)

ID	Mean	Median	Standard Deviation	Range
A3	4.625	5	1.7678	1-6
I1	5.75	7	1.9086	2-7
I2	5.625	6.5	1.9955	2-7
I3	4.875	6	2.5877	1-7
I4	4.75	5.5	2.252	1-7
I5	5.25	6	2.252	1-7
I6	5	5.5	2.2039	2-7
I7	5	5.5	2.1381	2-7
S1	5.875	6	1.3562	3-7
S2	6.25	6.5	1.0351	4-7
S3	6.125	6.5	1.3562	3-7
S4	5.875	7	1.6421	3-7
S5	5.75	6	1.2817	4-7
S6	5.375	6	1.8468	3-7
K1	5.625	6	1.685	2-7
K2	5.75	6.5	1.7525	2-7
K3	5.625	6	1.1877	3-7
K4	5.625	6	1.685	2-7
K5	5.5	5.5	1.6903	2-7
GF1	5	5	1.5119	2-7
GF2	4.5	4.5	0.9258	3-6
GF3	5.25	5.5	1.669	2-7
GF4	5.125	6	2.1002	1-7
GF5	4.75	4	1.7525	2-7
GF6	5.875	6.5	1.7269	2-7
GF7	4.25	5	2.0529	1-7
GF8	5.125	5.5	1.9594	1-7
GF9	5.625	6	1.598	2-7
GF10	5.5	6	1.7728	2-7
GF11	5.625	5.5	1.4079	3-7
GF12	4.5	5.5	2	1-6
GF13	5.375	6	1.598	3-7
GF14	4.75	5	1.669	2-7
GF15	4.75	5	1.9086	1-7
GF16	5.25	5.5	1.669	2-7

Table 3.1: Results of the questionnaire (N=8) | Raw response data can be viewed in Table A.1

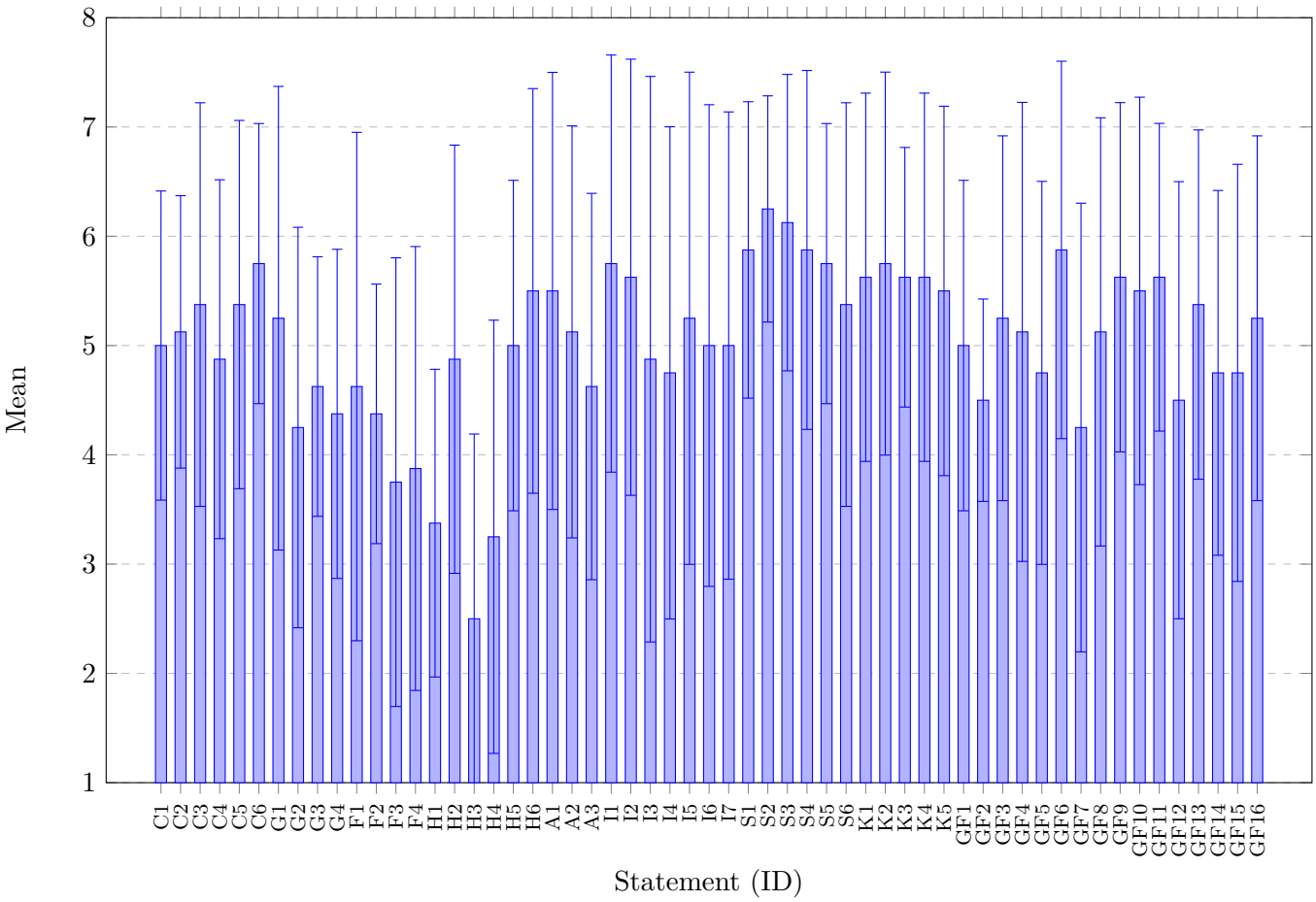

Figure 3.1: Mean and Standard Deviation of Responses for Each Statement (N=8) | Visualised data from Table 3.1.1 | Bars represent the mean values of the responses with markings indicating the standard deviation from this mean. | C=Concentration, G=Goal Clarity, F=Feedback, H=Challenge, A=Control/Autonomy, I=Immersion, S=Social Interaction, K=Knowledge Improvement, GF=Game Factors

	Mean	Median	Standard Deviation	Range
Stars	4.125	4.5	1.126	2-5

Table 3.2: Overall Rating of Siguru on a Scale from 1-5 (Star Rating)

3.1.2 Overall Rating

To receive an overall rating from players a simple star rating is filled out at the end of the questionnaire. This replaced the visual overall rating on a scale from 1-100 from the questionnaire basis by Fu et al. (2009). See Table 3.2 for the overall rating by the players.

$$\alpha_{total} = \frac{Q}{Q-1}(1 - \frac{\sum {\sigma^2}_{individual}}{{\sigma^2}_{total}})$$

$$\alpha_{total} = \frac{57}{57-1}(1 - \frac{177.142857142857}{5005.35714285714})$$

$$\alpha_{total} = 0.9818$$

Figure 3.2: Cronbach's alpha indicating validity and reliability | Q = No. of statements, ${\sigma^2}_{individual}$ = Variance of individual statements and ${\sigma^2}_{total}$ = Variance of all statements

$\alpha_C = 0.8356$
$\alpha_G = 0.8206$
$\alpha_F = 0.5721$
$\alpha_H = 0.6171$
$\alpha_A = 0.9544$
$\alpha_I = 0.9158$
$\alpha_S = 0.9423$
$\alpha_G = 0.9630$
$\alpha_{GF} = 0.9598$

Figure 3.3: Cronbach's alpha for individual dimensions (see Appendix for data) | Dimensions for which $\alpha < 0.8$ are marked red as they diverge from the original results by Fu et al. (2009)

3.1.3 Validity and Reliability

To measure whether the modification of the original questionnaire by Fu et al. (2009) invalidated or strongly reduced the validity or reliability which for the original results were $\alpha = 0.942$ for all 42 items and $\alpha > 0.8$ for each separate dimension (Fu et al., 2009) Cronbach's alpha was calculated in the same manner. The importance of Cronbach's alpha as a metric for measuring reliability of evaluation instruments is summarised by Nunnally and Bernstein, 1994:

> Cronbach's coefficient alpha (α) is perhaps the most important outcome, as it provides actual estimates of reliability. The α is basically the ratio of the sum of the covariances among the components of the linear combination (items), which estimates true variance, to the sum of all elements in the variance-covariance matrix of measures, which equals the observed variance.

The total result 3.2 indicates that the modified questionnaire/scale used has excellent internal consistency and reliability and did not compromise the original high score. However, for separate dimensions *Feedback (F)* and *Challenge (H)* the results differ

strongly from the original scale by Fu et al. (2009). All other dimensions clear the $\alpha_{dimension} > 0.8$ threshold determined by Fu et al. (2009) to be satisfactory. A potential factor that may compromise the reliability of the original dimensions is the fact that EGameFlow is intended for e-learning games rather than physical education games. Due to this dimensions such as *Feedback*, as will be addressed in the following Section 3.2, may be affected more strongly than others. Additional trials and measured playtests of Siguru with a vastly larger participant pool should be carried out in an effort to pinpoint the origin of this variance and to ensure the measured consistency is representative.

3.2 Interpretation & Discussion of Results

The results are measured on a Likert scale from 1 to 7 with each statement showing its mean, median, standard deviation and range to make the conclusions and interpretations from this data transparent. The goal is to understand the central tendency and variability of the participants' responses. This analysis will interpret this data and derive insights about the players' opinions.

3.2.1 Results with a negative trend

Starting by analysing mean values which indicate the average opinion of all participants for a statement the data in Figure 3.1 shows a positive reception of the pilot game test of Siguru since the majority of mean ratings lie above average with the threshold. This threshold between positive and negative trends being marked as the middle of the Likert scale at a value of 4.5. The statements' means below the 4.5 threshold are especially interesting as they mark areas that may need to be improved and are comprised of:

G2 Overall game goals were presented clearly

G4 Intermediate goals were presented clearly

F2 I receive immediate feedback on my actions

F3 I am notified of new events immediately

F4 I am notified of new tasks immediately

H1 The game provides "hints" in text that help me overcome the challenges

H3 The game provides video or audio auxiliaries that help me overcome the challenges

H4 The difficulty of challenges increase as my skills improved.

GF7 My in-game achievements felt uniquely mine.

Statements' medians exactly at the 4.5 threshold:

GF2 I was inspired by the overarching story or theme of the game.

GF12 I felt eager to progress and unlock new content in the game.

Taking the medians of statements with lower mean evaluations into consideration - for example *H3*: 2.5, *H1*: 3.5 or *F3*: 4 - reinforces the players' generally negative opinions of these statements. For many values the standard deviance and range are wide which indicates that players did not necessarily share the same opinions. Ergo the few values with low range and low standard deviance the evaluation can be seen as a very strong consensus. The results for the dimension *Goal Clarity (G)* show that players do not consider the goals to be very clear. This sentiment is reflected in Section 3.3 where

players give additional feedback about the scoring system and action cards. The values suggest that the instructions and presentation of how the game is played needs fine-tuning to ensure all players understand the goals of the game and how to play it.

The *Feedback (F)* statements received lower ratings with only one of the statements barely surpassing the middle of the scale. While it can be argued that a teammate being able to correctly guess the term provides immediate feedback on the effectiveness of communication, players do not to consider this as sufficient or immediate feedback. As Siguru is a physical card game it is challenging to introduce other feedback avenues that do not stem directly from the social interactions between players. While this result was anticipated in the original thesis (Thomas, 2024), future updates and features of Siguru may be able to address this topic.

Regarding *Challenge (H)* some mean values may come as a surprise as the extensive text-based explanations accessible by QR code on the respective card should provide all information needed. However, the observations in Section 3.3 do note that the players rarely took the time to scan the QR codes and read up on terminology. Instead the players opted to discuss with each other which can be considered a highly desirable alternate action. When neither of the teams were familiar with a concept they did fall back to the text-based resources provided on the cards. The game, as of finalisation of the original thesis (Thomas, 2024), does not provide any video or audio auxiliaries aside from the game designer being present during this play test to provide explanations and references to (video) snippets of pop culture in the card content. Raising this rating for a physical card game by introducing videos is a challenging aspect that could be solved by supplying the content on the game's homepage. However, observations should be made whether adding this auxiliary help lowers other ratings due to slowing the game down or similar consequences. The skill-based progression of challenge generally lies in the hands of the players that are free to add higher levels of cards with more difficult and advanced topics. However, the game does not offer a dynamic scaling of difficulty directly tied to player skill nor does it enforce any progression. An additional rule, for example: If a certain score is reached in the first match the team has to draw half of their hand from the next difficulty level, may raise this rating in future evaluations. Again, observations on how this change affects other factors should be made.

3.2.2 Results with neutral trend

Having discussed the lower scores, it is important to note that some of them result from limitations of the medium as a physical card game to be played in teams, e.g. offering dynamic skill scaling, notification and feedback systems and video/audio auxiliary materials. Vast changes in game structure will affect the evaluation outcome of other factors. Quite a few statements hover around a neutral rating, e.g. *G3*, *F2* and *GF2* which are in the middle of the scale. The medians for these values suggest that players have differing opinions resulting in a balanced overall evaluation.

3.2.3 Results with positive trend

The remaining majority of statements are very positive and align with most of the outlined expectations by Thomas, 2024. Especially statements related to *Social Interaction (S)* are worth highlighting as:

1. *S2 (I strongly collaborate with other classmates)*,
2. *S3 (The cooperation in the game is helpful to the learning)*,
3. *S4 (The game supports social interaction between players (chat, etc))*,
4. *GF9 (I enjoyed competing with other players in the game.)* and
5. *S1 (I feel cooperative toward other classmates)*

have the highest mean values indicating that players strongly agree with the statements.

The high medians and the positive range in ratings reinforce the observation that players enjoy the social interactions, working with each other and feel that it benefits the game. *Immersion (I)* and *Knowledge Improvement (K)* are the next two highly rated categories indicating a deep engagement with and enjoyment of the game which highlights an enjoyable and captivating game loop. While imparting new knowledge abouta topic is the ultimate goal of any educational material, the player's positive evaluation of Siguru draws attention to these additional dimensions.

3.3 Additional Feedback & Observations

The raw feedback is available in Appendix A.2.

3.3.1 Feedback

The common points of the additional feedback provided by players are:

- Sticking with the same hand of 7 term cards throughout the three rounds is criticised multiple times
- The rules for the action cards are unclear
- What each action card does is unclear
- The wish for a short description on the term cards is expressed
- The scoring system and instructions are not clear in the beginning

Many of these central points are reflected in the questionnaire data with the feedback providing some additional insight into why certain statements scored lower than one may expect. For example, the *Goal Clarity* evaluations could be attributed to the unclear scoring system, instructions and rules for the cards or the results of the statements regarding hints in the *Challenge* reflecting the wish for short descriptions on the cards instead of separated QR codes.

3.3.2 Observations

During the tournament play test observations and verbal feedback were documented. The list of general observations:

- The presentation explaining the game and its rules only was understood once the players saw the physical cards
- Many questions about how the action cards work arose from all teams/tables/-matches during the gameplay
- Verbal feedback that players should redraw cards between rounds was noted
- Reactions after the first round were expressively positive
- The opposing teams keep the other team in check regarding rules → self-regulating
- Across the tables variations in how the rules were understood appeared causing the tables to play different versions of the game
- Verbal feedback on being annoyed having to look up the terminology using the QR code was noted. Their teammate suggested that not knowing terms is supposed to cost time and act as punishment
- A team with strong knowledge about information security played the level 3 cards and gave verbal feedback saying it is a challenge
- The participants (cybersecurity students) expressed surprise that they learned new terminology previously unknown to them
- After the first team switches between tables had occurred confusion regarding the scoring system arose → the teams had been playing different rules and had to agree on a common denominator
- Unknown cards tend to be put aside in favour of making it through the remaining cards instead of looking them up
- Unknown cards were discussed after the round between teammates and with opponents instead
- After the tournament ended various verbal positive feedback about the game being fun, engaging and educational was noted. However, players suggested that the game introduction and rules still need polishing

A majority of these points are repeating the additional feedback 3.3.1 already provided. However, the observations give extra context surrounding the player feedback and indicate that a major portion of negative feedback stems from an initial misunderstanding of the rules. After familiarising themselves with the cards and the game mechanics the players quickly found their rhythm and offered improvement suggestions instead of direct critique which shows that playing the game cleared up most of the confusion.

Chapter 4

Closing

4.1 Case-study conclusion

In conclusion this case study has shown Siguru by Nohlberg (2024) to be a working example of a learning game that teaches information security concepts. The findings from the questionnaire, the analysis based on existing research on the topic of gamification of learning materials, the player feedback and the separate observations indicate that this physical card game excelled at fostering social interactions in a balance of cooperative and competitive play. Additionally, the game succeeded in teaching students of information security that can be considered subject-matter experts new terminology and concepts and provided a challenge for even the most advanced players with the more difficult cards according to observations 3.3.2. While the deck of cards already contains a broad spectrum of information security terminology the game mechanics and general framework can easily be ported to a different field of expertise by creating a new deck of cards. Alternatively, the base game can be expanded by simply adding cards and card content to the already expansive content.

The positive reactions by the players reinforce the game's perceived ability to make education about information security interactive, enjoyable and interesting while also enabling discussion about the topic between players in a casual learning environment. However, the case study did face limitations due to the small sample size caused by a small player pool and an even smaller pool of valid questionnaire samples because of which the game may require additional evaluation in the future to ensure that the data collected in this case study is universally reflected. Overall, Siguru, the educational card game has proven itself to be an effective tool and assistance to parties that teach information security concepts and spread awareness. It has potential for broader applications, expansion and refinement based on the insights gained from this case study.

Appendix A

Appendix

A.1 Questionnaire raw results

ID	Mean	Median	SD	Range	P1	P2	P3	P4	P5	P6	P7	P8
C1	5	5	1.4142	3-7	5	5	3	4	7	5	4	7
C2	5.125	5	1.2464	3-7	4	5	3	6	7	5	5	6
C3	5.375	6	1.8468	2-7	2	4	4	6	7	6	7	7
C4	4.875	5	1.6421	2-7	4	4	4	6	7	6	2	6
C5	5.375	6	1.685	2-7	2	4	5	6	6	7	6	7
C6	5.75	6	1.2817	4-7	4	6	5	6	7	7	7	4
G1	5.25	6	2.1213	2-7	2	6	6	5	7	7	7	2
G2	4.25	4.5	1.8323	2-7	2	5	5	4	2	6	7	3
G3	4.625	5	1.1877	3-6	3	4	5	5	6	6	5	3
G4	4.375	4	1.5059	3-7	3	3	5	4	7	6	4	3
F1	4.625	5.5	2.3261	1-7	3	2	6	5	7	7	1	6
F2	4.375	4	1.1877	3-7	4	4	4	4	7	4	3	5
F3	3.75	4	2.0529	1-7	4	1	1	3	5	5	7	4
F4	3.875	3.5	2.031	1-7	3	2	1	3	6	5	7	4
H1	3.375	3.5	1.4079	1-5	3	2	1	3	5	4	5	4
H2	4.875	5.5	1.9594	2-7	2	3	7	6	7	6	5	3
H3	2.5	2.5	1.6903	1-6	3	2	1	3	1	6	1	3
H4	3.25	3.5	1.9821	1-7	2	4	1	3	1	4	7	4
H5	5	5.5	1.5119	3-7	3	5	4	6	6	6	7	3
H6	5.5	6	1.8516	2-7	2	5	7	5	7	7	7	4
A1	5.5	6	2	1-7	1	5	6	7	7	7	6	5
A2	5.125	6	1.8851	2-7	2	4	7	7	6	6	6	3
A3	4.625	5	1.7678	1-6	1	5	5	6	6	5	6	3
I1	5.75	7	1.9086	2-7	2	5	4	7	7	7	7	7
I2	5.625	6.5	1.9955	2-7	2	6	3	7	7	6	7	7
I3	4.875	6	2.5877	1-7	1	6	1	7	7	6	7	4
I4	4.75	5.5	2.252	1-7	2	5	1	7	6	6	7	4
I5	5.25	6	2.252	1-7	1	5	3	7	7	7	7	5

Table A.1 (*continued*)

ID	Mean	Median	SD	Range	P1	P2	P3	P4	P5	P6	P7	P8
I6	5	5.5	2.2039	2-7	2	3	3	7	7	7	7	4
I7	5	5.5	2.1381	2-7	2	5	2	7	6	7	7	4
S1	5.875	6	1.3562	3-7	3	5	6	7	7	6	7	6
S2	6.25	6.5	1.0351	4-7	4	6	6	7	7	6	7	7
S3	6.125	6.5	1.3562	3-7	3	6	6	7	7	6	7	7
S4	5.875	7	1.6421	3-7	4	7	3	7	7	5	7	7
S5	5.75	6	1.2817	4-7	4	6	4	5	7	6	7	7
S6	5.375	6	1.8468	3-7	3	5	3	7	7	4	7	7
K1	5.625	6	1.685	2-7	2	6	5	7	7	5	7	6
K2	5.75	6.5	1.7525	2-7	2	5	7	7	7	6	7	5
K3	5.625	6	1.1877	3-7	3	6	6	7	6	6	6	5
K4	5.625	6	1.685	2-7	2	5	6	7	7	6	7	5
K5	5.5	5.5	1.6903	2-7	2	5	5	7	5	7	7	6
GF1	5	5	1.5119	2-7	2	4	5	6	7	6	5	5
GF2	4.5	4.5	0.9258	3-6	3	4	6	5	5	5	4	4
GF3	5.25	5.5	1.669	2-7	2	6	4	7	7	6	5	5
GF4	5.125	6	2.1002	1-7	1	6	6	7	6	7	3	5
GF5	4.75	4	1.7525	2-7	2	6	4	4	7	4	7	4
GF6	5.875	6.5	1.7269	2-7	2	7	6	7	7	5	7	6
GF7	4.25	5	2.0529	1-7	2	5	1	5	5	6	7	3
GF8	5.125	5.5	1.9594	1-7	1	6	5	7	5	7	6	4
GF9	5.625	6	1.598	2-7	2	6	6	6	7	7	5	6
GF10	5.5	6	1.7728	2-7	2	7	4	7	6	7	6	5
GF11	5.625	5.5	1.4079	3-7	3	5	5	7	6	7	7	5
GF12	4.5	5.5	2	1-6	1	4	6	5	2	6	6	6
GF13	5.375	6	1.598	3-7	3	3	5	7	6	6	7	6
GF14	4.75	5	1.669	2-7	2	5	4	7	6	6	3	5
GF15	4.75	5	1.9086	1-7	1	5	5	6	7	5	3	6
GF16	5.25	5.5	1.669	2-7	2	6	4	7	7	5	5	6
Overall	4.125	4.5	1.126	2-5	2	4	4	5	5	5	5	3

Table A.1: Raw results of the questionnaire (P#=Participant#; N=8)

A.2 Raw Additional Feedback

- "Always playing the 7 cards in every round feels kinda wrong"
- "A lot of attack names was unknown or was known as an other term"

- “The ”action cards“ are not self explaining, hardly got any of them right. The 7 Crads should not be used like this cause u get used to the cards and can easy guess them. Would be better to draw a new card for every guess”
- “Some cards have a matter of interpretation, which is funny but sometimes it can be frustrating because you can’t really argue while in game. The defence cards should be specified a bit more. Sometimes you can use them as beginning cards and defence cards which is kind of weird especially while playing the game. Some cards seem close to impossible to guess in the first round and later you have a keyword for certains cards which lets you instantly know the card. But all in all it is a really fun game to play. The mechanics are cool and i would play it again. I would even buy it if it is a little bit optimized. I really enjoyed playing the game and we had a lot of fun. Also i find the aspect of twisting the rules a little bit really interesting and new. As long as you have some groundrules it is nice to play and really intense. One final idea would be to flip a coin if both teams really can’t find a way to resolve a loophole so you have a decision.”
- “some attack cards that are ”defense“ need to be multi type ”defense / before“ to make sense attack cards need to be able to be used against one self no qr codes but short descriptions pls”
- “Some of the Action cards would benefit from clearer instructions on their effects (for example the shoulder surfing card). Redrawing the term cards every round would make the consecutive rounds more interesting since as it is you simply have to learn what cards your team has in the first round and cycle trough those in the following rounds.”
- “I liked it very much. The start was a bit difficult though. Because you don’t really get it how it works just by oral explaination (is like this for every game). For a better start playing 1 round, or even just one of the 3 little tounds with a more experienced player could make it better and more understandable. (I know we did not do this on purpose, just pointing it out)”
- “Scoring System was not clear in the beginning”

A.3 Adapted Questionnaire

Factor	**ID**	**Content**
Concentration	C1	Most of the gaming activities are related to the learning task
	C2	No distraction from the task is highlighted
	C3	Generally speaking, I can remain concentrated in the game
	C4	I am not distracted from tasks that the player should concentrate on
	C5	I am not burdened with tasks that seem unrelated
	C6	Workload in the game is adequate

Table A.2 (*continued*)

Factor	ID	Content
Goal Clarity	G1	Overall game goals were presented in the beginning of the game
	G2	Overall game goals were presented clearly
	G3	Intermediate goals were presented in the beginning of each scene
	G4	Intermediate goals were presented clearly
Feedback	F1	I receive feedback on my progress in the game
	F2	I receive immediate feedback on my actions
	F3	I am notified of new events immediately
	F4	I am notified of new tasks immediately
Challenge	H1	The game provides "hints" in text that help me overcome the challenges
	H2	The game provides "online support" that helps me overcome the challenges
	H3	The game provides video or audio auxiliaries that help me overcome the challenges
	H4	The difficulty of challenges increase as my skills improved
	H5	The game provides new challenges with an appropriate pacing
	H6	The game provides different levels of challenges that tailor to different players
Autonomy	A1	I feel a sense of control and impact over the game
	A2	I always know the next step in the game
	A3	I feel a sense of control over the game
Immersion	I1	I forget about time passing while playing the game
	I2	I become unaware of my surroundings while playing the game
	I3	I temporarily forget worries about everyday life while playing the game
	I4	I experience an altered sense of time
	I5	I can become involved in the game
	I6	I feel emotionally involved in the game
	I7	I feel viscerally involved in the game
Social Interaction	S1	I feel cooperative toward other classmates
	S2	I strongly collaborate with other classmates
	S3	The cooperation in the game is helpful to the learning

Table A.2 (*continued*)

Factor	ID	Content
	S4	The game supports social interaction between players (chat, etc)
	S5	The game supports communities within the game
	S6	The game supports communities outside the game
Knowledge Improvement	K1	The game increases my knowledge
	K2	I catch the basic ideas of the knowledge taught
	K3	I try to apply the knowledge in the game
	K4	The game motivates the player to integrate the knowledge taught
	K5	I want to know more about the knowledge taught
Game Factors	GF1	The game made me feel like my actions had a significant impact
	GF2	I was inspired by the overarching story or theme of the game
	GF3	I felt a sense of accomplishment after completing game tasks
	GF4	The game effectively tracked and rewarded my progress
	GF5	I had multiple ways to achieve my goals in the game
	GF6	The game allowed me to come up with creative solutions to problems
	GF7	My in-game achievements felt uniquely mine
	GF8	The game allowed me to customize my experience
	GF9	I enjoyed competing with other players in the game
	GF10	I felt connected to other players through shared game experiences
	GF11	The game created a sense of urgency that kept me engaged
	GF12	I felt eager to progress and unlock new content in the game
	GF13	Unexpected events in the game kept me engaged
	GF14	I was constantly curious about what would happen next in the game

Table A.2 (*continued*)

Factor	ID	Content
	GF15	I was driven to avoid making mistakes in the game
	GF16	The risk of losing in the game kept me focused

Table A.2: The modified EGameFlow questionnaire based on the scale by Fu et al. (2009) and on additional game drivers sourced from An (2020) and Chou (2015).

Participant No.	C1	C2	C3	C4	C5	C6	G1	G2	G3	G4	F1	F2	F3	F4
1	5	4	2	4	2	4	2	2	3	3	3	4	4	3
2	5	5	4	4	4	6	6	5	4	3	2	4	1	2
3	3	3	4	4	5	5	6	5	5	5	6	4	1	1
4	4	6	6	6	6	6	5	4	5	4	5	4	3	3
5	7	7	7	7	6	7	7	2	6	7	7	7	5	6
6	5	5	6	6	7	7	7	6	6	6	7	4	5	5
7	4	5	7	2	6	7	7	7	5	4	1	3	7	7
8	7	6	7	6	7	4	2	3	3	3	6	5	4	4
Total														
Variance	2	1.5536	3.4107	2.6964	2.839e	1.6429	4.5	3.3571	1.4107	2.2679	5.4107	1.4107	4.2143	4.125

Participant No.	H1	H2	H3	H4	H5	H6	A1	A2	A3	I1	I2	I3	I4	I5	I6	I7
1	3	2	3	2	3	2	1	2	1	2	2	1	2	1	2	2
2	2	3	2	4	5	5	5	4	5	5	6	6	5	5	3	5
3	1	7	1	1	4	7	6	7	5	4	3	1	1	3	3	2
4	3	6	3	3	6	5	7	7	6	7	7	7	7	7	7	7
5	5	7	1	1	6	7	7	6	6	7	7	7	6	7	7	6
6	4	6	6	4	6	7	7	6	5	7	6	6	6	7	7	7
7	5	5	1	7	7	7	6	6	6	7	7	7	7	7	7	7
8	4	3	3	4	3	4	5	3	3	7	7	4	4	5	4	4
Total																
Variance	1.9821	3.8393	2.8571	3.9286	2.2857	3.4286	4	3.5536	3.125	3.6429	3.9821	6.6964	5.0714	5.0714	4.8571	4.5714

Participant No.	S1	S2	S3	S4	S5	S6	K1	K2	K3	K4	K5
1	3	4	3	4	4	3	2	2	3	2	2
2	5	6	6	7	6	5	6	5	6	5	5
3	6	6	6	3	4	3	5	7	6	6	5
4	7	7	7	7	5	7	7	7	7	7	7
5	7	7	7	7	7	7	7	7	6	7	5
6	6	6	6	5	6	4	5	6	6	6	7
7	7	7	7	7	7	7	7	7	6	7	7
8	6	7	7	7	7	7	6	5	5	5	6
Total											
Variance	1.8393	1.0714	1.8393	2.6964	1.6429	3.4107	2.8393	3.0714	1.4107	2.8393	2.8571

Participant No.	GF1	GF2	GF3	GF4	GF5	GF6	GF7	GF8	GF9	GF10	GF11	GF12	GF13	GF14	GF15	GF16	Total
1	2	3	2	1	2	2	2	1	2	2	3	1	3	2	1	2	139
2	4	4	6	6	6	7	5	6	6	7	5	4	3	5	5	6	272
3	5	6	4	6	4	6	1	5	6	4	5	6	5	4	5	4	246
4	6	5	7	7	4	7	5	7	6	7	7	5	7	7	6	7	337
5	7	5	7	6	7	7	5	5	7	6	6	2	6	6	7	7	351
6	6	5	6	7	4	5	6	7	7	7	7	6	6	6	5	5	337
7	5	4	5	3	7	7	7	6	5	6	7	6	7	3	3	5	333
8	5	4	5	5	4	6	3	4	6	5	5	6	6	5	6	6	283
Total																	5005.3571
Variance	2.2857	0.8571	2.7857	4.4107	3.0714	2.9821	4.2143	3.8393	2.5536	3.1429	1.9821	4	2.5536	2.7857	3.6429	2.7857	177.1429

Table A.3: Variance of Questionnaire Results

ID	C	G	F	H	A	I	S	K	GF	Overall
α	0.8356	0.8206	0.5721	0.6170	0.9544	0.9158	0.9423	0.9630	0.9598	0.9818

Table A.4: Cronbach's Alpha (α) of questionnaire data

References

Literature

An, Y. (2020). Designing effective gamified learning experiences. *International Journal of Technology in Education*, *3*(2), 62–69. https://doi.org/10.46328/ijte.v3i2.27 (cit. on pp. 4, 5, 20).

Chou, Y.-k. (2015). Actionable gamification: Beyond points, badges, and leaderboards. https://uxmx.club/wp-content/uploads/2020/05/Actionable-Gamification-Full-Book.pdf (cit. on pp. 4, 5, 20).

Fu, F.-L., Su, R.-C., & Yu, S.-C. (2009). Egameflow: A scale to measure learners' enjoyment of e-learning games. *Computers and Education*, *52*(1), 101–112. https://doi.org/https://doi.org/10.1016/j.compedu.2008.07.004 (cit. on pp. 4, 5, 8–10, 20).

Furdu, I., Tomozei, C., & Kose, U. (2017). Pros and cons gamification and gaming in classroom. *Broad Research in Artificial Intelligence and Neuroscience*, *8*(2), 56–62. https://doi.org/10.48550/arXiv.1708.09337 (cit. on p. 4).

Jia, Y., Xu, B., Karanam, Y., & Voida, S. (2016). Personality-targeted gamification: A survey study on personality traits and motivational affordances. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, 2001–2013. https://doi.org/10.1145/2858036.2858515 (cit. on p. 1).

Nohlberg, M. (2024). Siguru: The cyber security card game. Retrieved July 6, 2024, from https://www.siguru.net (cit. on pp. 1–3, 14).

Nunnally, J. C., & Bernstein, I. H. (1994). https://books.google.at/books?id=_6R_f3G58JsC (cit. on p. 9).

PricewaterhouseCoopers. (n.d.). Game of threats. https://www.pwc.com/sg/en/services/consulting/risk-consulting/cybersecurity/game-of-threats.html (cit. on p. 1).

Thomas, I. (2024). *Gamifying social media security and privacy awareness materials* [Unpublished bachelor's thesis]. (Cit. on pp. 1, 11, 12).

Online sources

Anadea. (2018). *How gamification in the workplace impacts employee productivity.* https://medium.com/swlh/how-gamification-in-the-workplace-impacts-employee-productivity-a4e8add048e6 (cit. on p. 1).

van den Boer, P. (2013). *Introduction to gamification (whitepaper)*. https://www.slideshare.net/pvandenboer/whitepaper-introduction-to-gamification (cit. on p. 1).