PERSUADED: Fighting Social Engineering Attacks with a Serious Game

Dina Aladawy[1], Kristian Beckers[1,3], and Sebastian Pape[2,3]

[1] Technische Universität München (TUM), Institute of Informatics
Boltzmannstr. 3, 85748 Garching, Germany
[2] Goethe University Frankfurt, Faculty of Economics and Business Administration
Theodor-W.-Adorno-Platz 4, 60323 Frankfurt am Main, Germany
[3] Social Engineering Academy (SEA) GmbH
Eschersheimer Landstrasse 42, 60322 Frankfurt am Main, Germany

Abstract. Social engineering is the clever manipulation of the human element to acquire information assets. While technical security of most critical systems is high, the systems remain vulnerable to attacks from social engineers. The challenge in defeating social engineering is that it is a deceptive process that exploits human beings. Methods employed in social engineering do not differ much from those used to perform traditional fraud. This implies the applicability of defense mechanisms against the latter to the context of social engineering. Taking this problem into consideration, we designed a serious game that trains people against social engineering using defense mechanisms of social psychology. The results of our empirical evaluation of the game indicate that the game is able to raise awareness for social engineering in an entertaining way.

Keywords: security controls, social psychology, gamification

1 Introduction

Chris Hadnagy [9] defines social engineering as "Any act that influences a person to take an action that may or may not be in their best interest". Kevin Mitnick told in an interview the following about the relevance of social engineering: "The hacker is going to look at the weakest link in the security chain, [...] if they see it's your people – if you don't educate your people about social engineering and they're easy targets – then that's where the attacker is going to attack."[6] Mitnick's statement was made over a decade ago and is still of utmost importance today as several current studies confirm [15, 4].

In a previous work, we provided a mapping between social psychology and IT-security regarding Social Engineering defence [17]. In particular, we analysed social psychology methods of training against persuasion and mapped them to trainings in IT security. One identified gap is the lack of using *inoculation*, the repeated confrontation of people with a challenging situation in order to trigger an appropriate response. Our contribution in this work is filling the identified gap with a serious game called *Persuaded*.

Djaouti et al. [5] define serious games as follows “A serious game or applied game is a game designed for a primary purpose other than pure entertainment.”. We choose a serious game, because games recently built a reputation for getting employees of companies involved in security activities in an enjoyable and sustainable way. Williams et al. [20] introduced the protection poker game to prioritise risks in software engineering projects. Shostack [18] from Microsoft presented his Elevation of Privileges card game to practice threat analysis with software engineers. Furthermore, games are used as part of security awareness campaigns [7] and particularly as a part of social engineering threat analysis [1].

Our contribution Persuaded has inoculation incorporated into the core game mechanics to trigger resistance to social engineering attacks through exposing people to realistic attack scenarios. We designed our serious game to achieve the following goals: (a) increasing awareness of social engineering, (b) training resistance to persuasion and (c) addressing the general population. In order to provide the validity of the attack scenarios, we took all of them from scientific publications. The game enables employees to learn about social engineering, while practicing simultaneously. This immediate application of learned knowledge has proven to have lasting effects [8].

The game works as follows. Employees get confronted with a possible social engineering threat and have to select a defense mechanism. This defense mechanism is a pattern of behaviour ensuring a secure outcome. For example, an employee gets a phishing mail and is asked to open its attachment. Afterwards the player selects a countermeasure: "Do not open the email and inform the information security department immediately". The player gets immediate feedback whether the chosen defense is correct. In particular, the offered defenses can be part from a company’s security policy. Non surprisingly, Soomro et al. found that development and execution of information security policy had a significant impact on the quality of management of information security [19]. Earlier, Pahnila et al. already concluded that appraisal and facilitating conditions have significant impact on attitude towards complying with the security policy while sanctions and awards do not have a significant effect on the intention to comply [14]. Thus, enabling employees to become familiar with the security policy in a playful way contributes to the holistic security of the company.

The remainder of the paper is organised as follows: We start with an overview of related work (Sect. 2) and a description of our game (Sect. 3). In the next sections we describe the study and its results. We end with a discussion of the results, threats to validity and the conclusion.

2 Related Work

As security is usually a secondary task, computer security training has often been perceived to be an uninteresting enforcement to users and managers. The approach of developing serious games has therefore been adopted to provide knowledge and training in that field.

CyberCIEGE is a role playing video game, where a player acts as an information security decision maker in an enterprise. Players' main responsibilities are to minimize the risk to the enterprise while allowing users to accomplish their goals. Similar to Persuaded, the game offers a simulation of the reality particularly portraying the need to maintain the balance between productivity and security. As decision makers, players get to make choices concerning users (i.e. How extensive will background checks be?), computers (i.e. How will computers be networked?) and physical security (i.e. Who is allowed to enter a zone?) while monitoring the consequences of their choices. When compared to Persuaded, we recognized CyberCIEGE offered several advantages common to those offered by Persuaded. For instance, players are in a defensive mode and they get to make decisions and experience their consequences. CyberCIEGE even incorporates assets and resources in the game, which is a missing element in Persuaded. On the other hand, the game requires longer time to learn and to play [10].

PlayingSafe is a serious game in the domain of social engineering. It consists of multiple choice questions which are wrapped in typical mechanics of a board game. Since questions provided are exclusive to social engineering, the game is very similar to ours. The main difference lies however in the focus in the topic of social engineering. PlayingSafe asks questions in the fields of Phishing, advanced fee fraud, spam and others, being a category that covers less common attacks. Our game on the other hand covers a broader field without offering depth in each topic. Additionally, our game incorporates strategy favouring the entertainment element, in order to enhance the game experience the game provides [12].

SEAG is a serious game designed to raise awareness of social engineering. The game utilizes levels that tackle different cognitive aspects and hence provide an effective learning experience. The first level consists of quiz-like questions to build a knowledge base for the players. The second level is a match game where players have to match social engineering terms with respective pictures. Finally, the players are presented with real life scenarios to analyse pertaining to threat. This simulation of real life application of the learnt lesson should test players ability to detect attacks- an approach very similar to inoculation [13].

3 Game Description

To fill the gap, identified by Schaab et al. [17], we designed a game that does not only provide knowledge, but rather trains people by implementing theories from social psychology on the resistance to persuasion. In this section, we give a brief overview of key design decisions, their rationale and our goals (cf. Fig. 1).

Game requirements: We refined our goals and report them in the following categorised by key areas of game design.
Ease to learn: A low level of complexity allows to learn the game more easily, and thus is more attractive to novices in game play.
Ease to play: To be easily integrated into the players' daily routine, the game should have a minimum of necessary preparations and a short play time. Given online games require less preparation than tabletop games, it should be online.

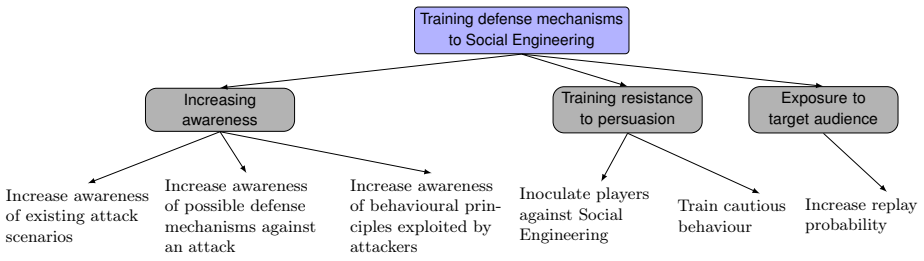

Fig. 1. Definition of goals for the game

Replay value: The replay value depends to a large amount on the ease to learn and play of the game. In order to maintain the appeal to expert players as well, game mechanics should provide a substantial entertainment element along with long term motivation and challenging the players. As multi-player games depend on the availability of other players, a single player game is preferred.
Player's role: In order to inoculate players against social engineering, they have to be in the position of an attack receiver.
Textual Content: Since our awareness goals cater for presenting attack/defense scenarios, the game design should support the presentation of textual content.

Game mechanics: In order to create a single player game with easy rules and low complexity, we decided to aim for a patience and solitaire game approach [11] instead of e.g. involving machine learning approaches [3] which would tend to result in a game with multi-player feeling. Thus, the player may choose between playing cards from his/her hand or draw the next card from the deck. As known from patience games, the deck is shuffled automatically for each game.

Types of cards and card functionalities: Four types of cards were chosen.

1. *Attack* cards include attack scenarios in textual form.
2. *Defense* cards describe a pattern of behaviour that protects the player against an exploitation attempt. A defense card exists for each attack card.
3. *See The Future* cards allow the player to take a peek on the three upper cards in the card deck.
4. *Skip turn* cards allow the player to take the upper card of the deck and put it below the deck.

Mechanics and rules: A turn in Persuaded consists of the following rules:

1. Play an action card or draw a card from the deck.
2. If you draw any card that is NOT an Attack, the turn is over. Put the card to your hand cards.
3. If you draw an Attack card, you *have to* play a Defense card. The correct (wrong) defense gains you 10 (-5) points. The Defense card is only discarded if you had a correct match. Otherwise it's put back in the deck.
4. If you draw an Attack card and don't have *any* Defense card in your hand, you lose one heart (life). If you lost all three hearts the game is over.
5. The game is won if the deck is empty and is lost if the player loses all 3 lives before finishing the deck.

These mechanics have several consequences. Drawing an Attack card *forces* the player to play a Defense card. Thus, even if a player notices he has no matching defense, he has to burn a defense card. This was introduced to further encourage cautiousness when drawing cards from the deck. The player needs to use *See the future* cards to have a peek on the pile and then postpone attacks if he does not have a matching defense by playing a *Skip turn* card. This also forces the player to match upcoming attacks and defenses in hand before drawing form the pile.

Long term motivation: As known from patience games, the deck is shuffled automatically for each game. This causes each game to be different from the game(s) before. Thus, the player needs to come up with different moves to win the game and can not simply try until he/she finds the 'optimal solution'. Additionally, the introduced randomness, causes Attack cards to appear before their respective Defense cards in the deck. Therefore – if action cards are not distributed accordingly – this may lead to situations where the player simply has to guess what might be the 'best next move'. The idea behind this rationale is that not only has the player to learn how to make best use of "See the future"- and "Skip turn"-cards, but also needs to have some luck in order to achieve the best possible score. We balanced it in a way, that it is always possible to win, but might not be possible to get the maximum score.

Game content: In order to provide the knowledge needed to increase players' awareness, scenarios of attacks and their respective defenses were incorporated in the game. We selected eight attack scenarios that represent different social engineering attack types, namely Baiting, Phishing, Tailgating, Mail attachment, physical and virtual Impersonation, Voice of Authority and Popup Windows. The attacks were inspired by a card game for eliciting security requirements [2]. Defense cards, on the other hand, confronted us with challenges, as it is not very intuitive to act against behavioural principles, which is exactly the element exploited by social engineering. We identified explicit defenses encouraged from best practice by security departments in different companies. Initially, defenses were meant to be generic and applicable for several attack scenarios. However, resulting from our selection of proposed scenarios, we noticed, that all had similar generic defenses, i.e. to verify the source or the person. Hence, we decided to incorporate one-to-one matches thereby providing eight specific Defense cards.

Game interface design: In confirmation with Don Normann's Design principles [16] for user interface design, we opted for an intuitive user interface that adheres to the needs of novices as well as experts in game play. The proposed design was further tested and adapted according to the feedback we received during the piloting phase. We used different colors for each type of the *cards* (see Fig. 2). For the attack/defense scenarios, we kept the text as short as possible and divided the content in up to three bullet points. Action cards consist of graphics that reflect their functionality, attack and defense cards have titles summarizing their content. However, titles of matching pairs are not the same. This design decision was intentionally incorporated, in order to assure that players have to read the cards' contents. The *Game Setting* (see Fig. 3) was designed to be both intuitive and informative.

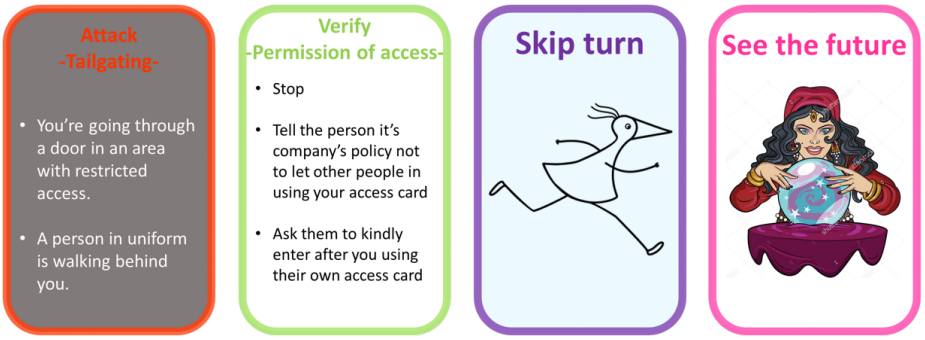

Fig. 2. From left to right: Attack card, Defense card, Skip card, See the future card

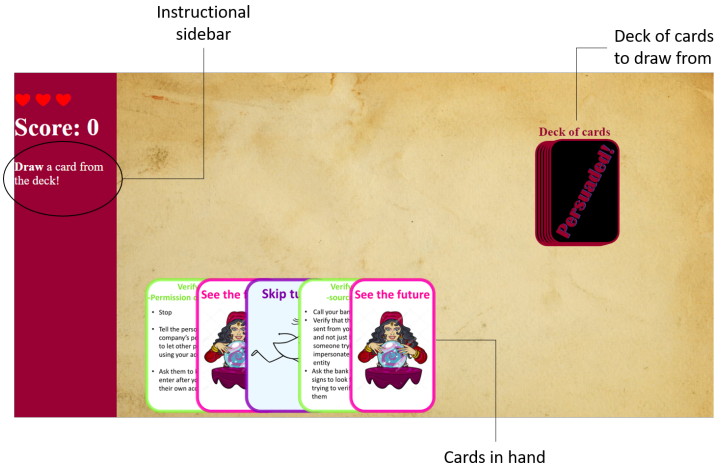

Fig. 3. Initial game setting

Cards in hand: The overlapping display of the current cards in hand simulates the holding of cards in real life (cf. Normann's mapping principle). When a player moves the mouse over a card, this card is emphasized by moving the other cards to the left and right to allow the player a complete view of the card. This enhances the player's experience while maintaining readability of the content.

Scoring: As score and lives function as a reward and punishment system, it is important to make sure, they capture players' attention when they change. Therefore, we decided to reflect modifications of scores and lives using dynamic feedback. In addition to using coloured terms such as "Defended", "Wrong match" and "Persuaded", the decrease or increase of score and lives is at the top left.

We show the game in detail in our Video Tutorial for Persuaded[4]. Furthermore, we stored the data of our experiment and an extensive technical report online[5].

[4] https://youtu.be/UWhc1e6ngd0
[5] https://sites.google.com/site/researchpersuaded/

4 Study Design

Prior to conducting the case study, Persuaded has been evaluated through several rounds during the design and the development phase. First of all, the scenarios were tested for suitability of attacks and defenses, in addition to the ease of understanding of the presented content. Following this, the game's functionality and mechanics were tested during a piloting phase. The participants for the pilot tests were very heterogeneous. Play tests and semi-structured interviews were conducted with 3 security experts, a psychology expert, a games engineering student, an informatics student and a philosophy student. Feedback provided in this phase, was largely incorporated in the design and the implementation. The content was reviewed by 2 security experts and 2 informatics students. The scenarios' text was reviewed by a student in Translation Studies.

4.1 Preparation and Collection of Data

The flow of a session with a subject consisted of the following steps:

1. Answer the pre-questionnaire.
2. Watch the game tutorial as many times as you need.
3. Ask questions about the game rules.
4. Play the game.
5. Answer the post-questionnaire.

We employed first and second degree methods for our data collection. Before the session started, subjects were encouraged to provide feedback throughout the session. Many subjects took this into account and offered valuable feedback on the questionnaire, the game and the tutorial. Some subjects even played the game in a think aloud mode, which turned out to be very useful feedback. Furthermore, second degree data was collected during the game play to evaluate to what extent the game adheres to requirements specified in prior sections. We logged all decisions made during the game, making it possible to replicate the entire round. In addition, the time to play as well as the final score and number of lives left was collected. This enabled us to analyse the effects of our random factor on the entire game experience.

Pre- and post questionnaires The effect of inoculation can be measured by observing peoples' reactions to stronger persuasive attacks as the ones they were inoculated with. This implied that we have to present players with stronger scenarios of social engineering after the game in order to be able to derive whether it was effective or not. This however, was not enough as an effect measurement as we were not aware, whether people were vulnerable before the game at all or not. Hence, we decided to conduct questionnaires before and after the game was played. The questionnaires presented social engineering scenarios as single choice questions, where players had to choose one of the given behaviours as a reaction to the given situation. The same scenarios with the same reactions were presented both in the pre-and post-questionnaire. This was intentionally done in order to be able to measure effects of the game as change of answers. In

addition to the situations presented in the pre-questionnaire, demographic data was collected to draw conclusions for different types of people given our exposure goals. Moreover, data concerning technical background was collected which might be relevant to scenarios such as *Phishing* and *Popup Window* as well as malicious *Mail Attachments*. Lastly, items were used to measure background knowledge of social engineering and to measure the subjective perception of vulnerability in order to have an indicator of optimism bias. Players were also asked to indicate at which point they understood the game to measure the learning curve and the effectiveness of the tutorial and whether they would play this game again or not.

4.2 Data

The equation introduced to evaluate the questionnaires was:

Learning outcome = $\sum$ security-aware behaviour in post-questionnaire $-\sum$ security-aware behaviour in pre-questionnaire

For quantitatively evaluating the players' decisions throughout the game, we relied on the following data.

Matching of attacks and defenses We used a half automatic analysis process to measure the number of attacks that were correctly defended as well as the number of burned defense cards. This data maps the understandability of the content of the cards. Moreover, as we are not game designers we decided to use them as an indicator of the impact of certain game elements such as the randomness of the cards' order and the variability of Attack and Defense cards.

Usage of action cards We also collected data concerning the number of cards that were foreseen and the number of cards that were unknowingly drawn. This information was not only used to evaluate the game flow, given that Flow cards are key elements of the winning strategy. They were also used as an indicator of players' risk behaviour and alertness during the game play.

Reward and punishment system Lastly we collected the score data as well as the number of lives left. This information was employed to test whether our reward and punishment systems are effective or whether they are influenced by the random element in the game.

5 Results

The study was conducted with 21 participants including 9 female and 12 male participants. The age ranged from 19 to 35 years. Given our exposure goals, we sampled subjects with different backgrounds regarding their studies. 16 of the participants indicated they are university students, while 5 are currently pursuing an academic career. We disregard one participant's results. These were invalid due to changes of the content of the questionnaires. In contrast to the variation in age and occupation, our sample is very homogeneous in technical background. It is important to mention that at this point, we only consider technical background in relation to how often the computer is used, which is sufficient for understanding the game content. Answering this question, 95%

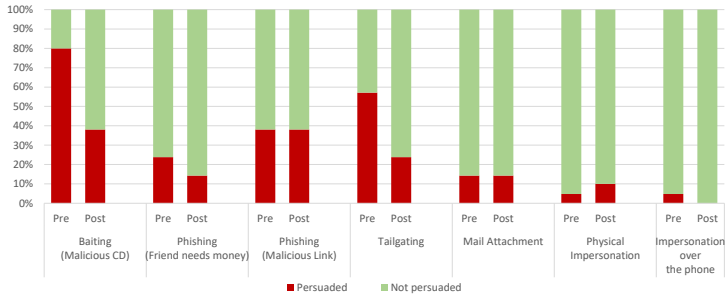

Fig. 4. Number of reactions reflecting falling for the attack in comparison to security aware reactions before and after the game

indicated they use their computers daily while 43% use it daily for job related matters.

5.1 Results relevant to inoculation

Our game is an implementation of inoculation against social engineering. Its effectiveness as a training method was evaluated using pre- and post-questionnaires in addition to several metrics.

Reactions to situations Participants were given social engineering scenarios and asked to choose one reaction, they were most likely to adopt if confronted with such a scenario. The questions proposed three answers with one mapping the security-aware behaviour when encountering a potential threat and the other two options reflecting extreme reactions. The first extreme is a paranoid reaction, whereas the second reflects falling victim for the attack. Results from the pre-questionnaire show that in 5 of the 7 scenarios the majority of the participants would have behaved in a manner that would not endanger them. In the other two scenarios, a high number of subjects would have fallen for the attack.

The results of the post-questionnaire show significant differences. For the *Tailgating* scenario which describes the situation of meeting a strange lady who is locked out of the house building and whether a person should verify her identity before letting her in or not, the number of participants indicating they would behave in a security-aware manner rises from 43% to 76%. Nevertheless, the *Baiting* attack which questions whether free handed CDs from street musicians should be scanned or not, remains the one scenario where the reaction indicating falling for the attack is the one chosen the most. For the *Phishing (Malicious Link)* and *Mail attachment* attack, the numbers do not show significant change. For the remaining scenarios only slight changes are noticeable, once even favouring the rise of number of participants who would fall for the attack as it is the case in the *Physical impersonation* attack. Figure 4 shows an overview of the change in responses triggered by our game. Given inoculation relies on repeatedly confronting individuals with mild persuasive attacks, we also measured the number of times players read an attack card in the game which indicates that each attack is read 1.5 times in average.

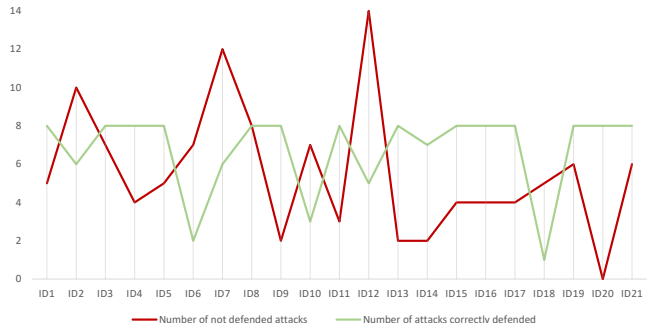

Fig. 5. Defended vs. Not defended attacks for all participants

5.2 Reward/Punishment system

The maximally achievable score is 80 points if the player did not make any wrong match or if the player did a wrong match at the beginning of the game when the score was still 0. Only one player was able to score 80 points and 2 players could score 75 with an average score of 51 points. The majority of players achieved a score of 65 points. Considering the lives maintained in the game, 15 players were able to finish the deck maintaining at least one heart while 6 others lost the game before finishing the deck due to losing their lives.

5.3 Time to play

The time needed to play the game ranged from 02:53 minutes to 16:03 with an average of 08:09 minutes. We further differentiate the time to play needed to win the game by finishing the cards in deck and the time to play for lost games. The range measured for games that were won through finishing the deck lies between 05:05 and 16:03 minutes with an average of 8:33 minutes.

5.4 Matching of cards

For 15 of the participants, the number of successfully defended attacks is higher or equal to the number of not defended ones. The latter further includes attacks that were drawn without having defense cards in hand. Not defended attacks can be further categorised in mismatched attacks (75% of not defended attacks) and attacks that were drawn without having defense cards in the hand (25% of not defended attacks).
We further distinguish mismatched attacks in those, where the player had the matching defense in hand, as in the player is accountable for the mismatch and those, where the player was forced to play a Defense card. For all participants the number of burned defenses is higher than the number of *truly* mismatched attacks. Moreover, 66% of the participants did not even once mismatch an attack while the matching Defense card is in hand.

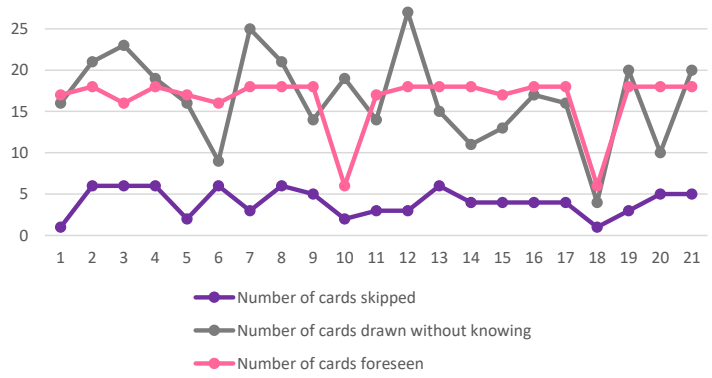

Fig. 6. Overview of skipped cards, cards foreseen and cards drawn blindly

5.5 Action cards

The key strategy to winning the game is to use *See The Future* cards and then avoid attacks, whose defenses are not in hand, with *Skip* cards. This is why we analysed players' usage of action cards considering players can see a maximum of 18 cards before drawing them and skip a maximum of 6 cards.

We further analysed players risk behaviour according to when cards are drawn blindly despite having a *See The Future* card. Our results show that a total of 16 players have drawn at least one card without seeing it, having a *See The Future* card in the hand. An average of 2.6 cards were drawn blindly despite having the chance to foresee them. More importantly, however, is the number of cards drawn blindly despite having both a *See The Future* card and a *Skip* card. This card combination would have offered the chance to knowingly avoid drawing that card. This was done by 10 of the participants with an average of 1.6 cards drawn blindly despite having the chance to knowingly avoid them.

5.6 Learning curve and Replay value

Finally, subjects were asked to indicate whether they would play this game again. 17 participants (81%) expressed that they would play this game again while the remaining 4 participants claimed they would not. When asked about the understandability of the game mechanics, 14 people (67%) mentioned they understood the game right away (following only the tutorial) while the remaining 7 participants needed some turns to fully understand how the game works.

6 Discussion

Through the conduction of interviews with the players, we could collect feedback that is of value to future work. More importantly, the feedback showed potential threats to the validity of the data collected in the questionnaires.

6.1 Feedback on pre-questionnaire:

The pre-questionnaire included social engineering scenarios, where players had to chose a reaction. Particularly, the baiting scenario, where street musicians would intentionally offer malicious CDs, was perceived by two participants to be an "interesting new attack, [they] have never thought of". The tailgating attack was stated to be relevant to one of the participants. Particularly these two attacks were the only ones in the pre-questionnaire, where most participants chose the reaction, that would favour the attackers intentions. We conclude, that these attacks were new to the participants choosing that reaction. This is backed up by the interview comments in addition to the results of the final question in the post-questionnaire where seven participants indicated the *Tailgating* attack was new to them while four indicated the same for the *Baiting* attack. Improvement suggestions were to incorporate an "other" option as a possible reaction to the situations and to collect data on the used operating system, given it implies a certain security level provided by the technology alone.

6.2 Feedback on the game mechanics:

We received extensive feedback on our randomness factor of our game.

General Perception: The general perceptions during game play provided feedback that conforms with our design goals. One participant stated, that for them the game simulates the reality. The player further explained, that in real life, it is rather difficult to expect social engineering attacks and always be ready for them, which they found was mapped through the random factor. Furthermore, the player mentioned that usually even the most cautious people might fall victim for social engineering again supporting the vulnerability in the game, where players are not able to defend themselves when drawing attacks before their respective defenses. Another participant provided feedback on the challenge level in the game saying that "one has to think". For this player, the game was also "easy to understand", reflecting the modesty of the trade-off between those two conflicting elements. Finally, the player emphasised the importance of the game being single player for the replay value, saying that he can "play the game another 3 times just right now". This data conforms with the data collected on replay in the post-questionnaire underlining the high potential for replay of the game.

Understanding of the game mechanics: We opted for ease of learning, realised by simple mechanics, a detailed tutorial and an intuitive game interface design. Several questions asked during game play, however, indicate otherwise. Misconceptions and uncertainty were particularly common regarding the functionality and usage of action cards. Examples for questions, we received concerning action cards are: *What does a "See the future" card really do?*, *What does a "Skip" card really do?*, *How many cards are skipped by playing a "Skip" card?*, *Will skipped cards be added in the deck?*. We cannot determine, whether these questions were asked due to lack of understanding or rather to confirm prior understanding of the functionality. However this data explains the relatively small numbers of wasted *See The Future* cards and *Skip* cards, which were played

without having foreseen what was being skipped. We assume the wrong usage of the action cards happened at the beginning of the game, as four players have indicated, they needed some turns to fully grasp the game mechanics. Five players asked for the number of cards in the deck. We assume, this was asked in order to develop certain strategies rather than to indicated extensive length of the game duration, which is further supported by our measurement of game duration being 09:45 minutes in average.

Card content: The serious element of Persuaded lies in the content of the cards. This is why, it is important to monitor whether cards are read in detail or not. Four players indicated after the first couple turns, that they have not been reading the cards' contents, while two others attempted to match the titles of attack and defense cards in the beginning. Still, all six players started reading the content of the cards after a couple of turns. We assume this was motivated by the punishments they received for wrong matches. This data is particularly relevant to data collected on the number of mismatched cards while having the correct defense in hand. Given this only happened to an average of 0.52 cards, we build the assumption that *truly* mismatched cards were rather a result of not reading the cards than a result of the complexity of the content. Still, one player further indicated, that the match between attacks and defenses was not always clear. This was however intentionally incorporated in the design, as we wanted players to reflect about the scenarios and the defenses instead of recognizing the matches from the cards. Furthermore, one player suggested, it would be better to see all the cards before playing the game to create a mental scheme of matching cards. Thereby, players can solely focus on training the strategy during the game and would strengthen the mental scheme by recalling matches between cards.

Randomness in the game: The randomness of the game was not very welcomed by the participants. Although one player indicated, it provided a simulation of real life, four other players perceived themselves to have no control in the game with two players evaluating the game as unfair. An important aspect influenced by the randomness is replay value. Replay value is usually supported by the probability of the player to excel in the game play. Having a random factor largely limits improvements in the game as players' decisions are only partially relevant for the game results. This was further confirmed by two players, who said they would only play the game again, if they could get better at it.

7 Threats to validity

We discuss potential threats to the validity according to Wohlin [21].

Construct validity Questions in the post-questionnaire are supposed to indicate probable reactions of the participants to given scenarios. There is, however, a possibility that participants remember their answers to the same questions during the pre- questionnaire. However, if players are aware, the game has educational purpose, this might lead to a conscious choice of the correct answers to indicate having understood the content. In addition, several metrics were derived from players' decisions during the game session as explained in the

previous section. This data is however subject to effects of concentration and motivation during the game. Moreover the results assume that the functionality of the game and the different cards is understood at the beginning of the game in contrast to the feedback received on the learning curve.

Internal validity We measure the learning outcome as the difference between the sum of correctly answered questions in the pre- and post-questionnaire. We cannot determine whether players are inoculated by the scenarios of the game or by the scenarios mentioned in the pre-questionnaire as these also reveal persuasive arguments used by social engineers. This effect of an inoculation at an early point is attempted to be overcome by hiding the subject of the study from participants until the questions of the pre-questionnaire are answered.

External validity We conducted the case study with a heterogeneous population regarding their educational background and could identify acceptability of the game even for subjects without prior knowledge in security or social engineering. However, our results regarding the effectiveness and the learning outcome of the game are to be considered taking the random factor of the game and other threats to validity into account.

8 Conclusion

We designed, implemented and evaluated a serious game for training social engineering defense mechanisms, called "Persuaded". Several goals were specified and refined to achieve the serious purpose of the game: *Increase awareness:* of attack scenarios, defense mechanisms and exploited behavioural principles. *Train resistance to persuasion* by inoculation against social engineering and to train cautious behaviour. Finally, to cater for *exposure to the general population* through increasing replay probability and ease of understanding of the social engineering threat. Results of our case study indicate great potential for the application of social psychology defense mechanisms to social engineering. Our serious game offers a tool for monitoring decision making processes and risk-taking behaviour. More importantly, it was successful at raising awareness to new attack scenarios in an entertaining way such that people would enjoy learning about social engineering and how they can defend themselves against it.

Acknowledgements

This research has been partially supported by the Federal Ministry of Education and Research Germany (BMBF) with project grant number 16KIS0240.

References

1. K. Beckers and S. Pape. A Serious Game for Eliciting Social Engineering Security Requirements. In *Proceedings of the 24th IEEE International Conference on Requirements Engineering*, RE '16, pages 16–25. IEEE Computer Society, 2016.

2. K. Beckers, S. Pape, and V. Fries. Hatch: Hack and trick capricious humans - a serious game on social engineering. In *Proceedings of British HCI 2016*, pages 1–3. ACM, 2016.
3. M. Bowling, J. Fürnkranz, T. Graepel, and R. Musick. Machine learning and games. *Machine Learning*, 63(3):211–215, 2006.
4. Dimensional Research. The Risk of Social Engineering on Information Security: A Survey of IT Profesionals, 2011. http://docplayer.net/11092603-The-risk-of-social-engineering-on-information-security.html.
5. D. Djaouti, J. Alvarez, and J.-P. Jessel. Classifying serious games: the g/p/s model. *Handbook of research on improving learning and motivation through educational games: Multidisciplinary approaches*, pages 118–136, 2011.
6. ENISA. Social engineering: Exploiting the weakest links. Whitepaper, https://www.enisa.europa.eu/publications/archive/social-engineering, 10 08.
7. M. Gondree, Z. N. J. Peterson, and T. Denning. Security through play. *IEEE Security and Privacy*, 11(3):64–67, 2013.
8. F. L. Greitzer, O. A. Kuchar, and K. Huston. Cognitive science implications for enhancing training effectiveness in a serious gaming context. *J. Educ. Resour. Comput.*, 7(3), 2007.
9. C. Hadnagy. *Social Engineering: The Art of Human Hacking*. Wiley, 2010.
10. C. E. Irvine, M. F. Thompson, and K. Allen. Cyberciege: gaming for information assurance. *IEEE Security & Privacy*, 3(3):61–64, 2005.
11. A. H. Morehead. *The Complete Book of Solitaire and Patience Games*. Read Books Ltd, 2014.
12. M. Newbould and S. Furnell. Playing safe: A prototype game for raising awareness of social engineering. In *Australian Information Security Management Conference*, page 4, 2009.
13. A.-S. T. Olanrewaju and N. H. Zakaria. Social engineering awareness game (seag): An empirical evaluation of using game towards improving information security awareness. In *Proceedings of the 5th International Conference on Computing and Informatics, ICOCI 2015*, 2015. (Accessed on 10/16/2016).
14. S. Pahnila, M. Siponen, and A. Mahmood. Employees' behavior towards IS security policy compliance. In *System sciences, 2007. HICSS 2007. 40Th annual hawaii international conference on*, pages 156b–156b. IEEE, 2007.
15. PWC. Information Security Breaches Survey 2016, 2016. https://www.pwc.be/en/documents/media-centre/publications/2016/informationsecuritybreachessurvey2016.pdf.
16. Y. Rogers, H. Sharp, J. Preece, and M. Tepper. Interaction design: beyond human-computer interaction. *netWorker:The Craft of Network Computing*, 11(4):34, 2007.
17. P. Schaab, K. Beckers, and S. Pape. Social engineering defence mechanisms and counteracting training strategies. *Information and Computer Security*, 25(2):206–222, 2017.
18. A. Shostack. *Threat Modeling: Designing for Security*. John Wiley & Sons Inc., 1st edition, 2014.
19. Z. A. Soomro, M. H. Shah, and J. Ahmed. Information security management needs more holistic approach: A literature review. *International Journal of Information Management*, 36(2):215 – 225, 2016.
20. L. Williams, A. Meneely, and G. Shipley. Protection poker: The new software security "game". *Security Privacy, IEEE*, 8(3):14–20, May 2010.
21. C. Wohlin, A. von Mayrhauser, P. Runeson, M. Höst, M. Ohlsson, B. Regnell, and A. Wesslén. *Experimentation in Software Engineering: An Introduction*. International Series in Software Engineering. Springer US, 2012.