Gamification of Cyber Security Awareness Training for Phishing against University Students

Jelle Nijland
University of Twente
P.O. Box 217, 7500AE Enschede
The Netherlands
j.g.w.nijland@student.utwente.nl

ABSTRACT

Users are the main source of Cyber Security breaches. However, Cyber Security Awareness training is viewed as useless and uninteresting by the users. These users feel as if the training is a secondary task, an obstruction, or a distraction from their primary work. This apathy poses a risk to organizations, as Cyber Security breaches cost businesses combined billions a year. Gamification can provide the solution by giving an engaging and interactive alternative to these mandated training sessions. A vulnerable subset of these users is university students. Research suggests these students are the most likely to be fooled by a phishing attack. Although there are already many Serious Games available, this paper focusses on designing an interesting and engaging gamified method of training. This game was created specific to university students. The novelty of this research can be found in its form and competitiveness. Finally, this training was evaluated against non-trained users. The analysis of the results of the survey adds to the scientific body of knowledge on phishing prevention training. Additionally, the Serious Game developed to test this approach can be added to the collection of games out there, narrowing the field of missing games based on phishing.

Keywords

Cyber Security Awareness, Gamification, Phishing, University students

1. INTRODUCTION

As the scale of cybercrime is on the rise, so is the cost [16]. It has been predicted by Morrow et al. in 2021 from Juniper Research, that cybercrime will cost businesses more than 5 trillion dollars by 2024. Unfortunately, Purkait et al., 2012 proposed that humans are the weakest element in Information Security. Researchers with IBM X-Force showed that phishing was the leading attack vector in one-third of all cases [23]. Lance et al., (2005) defines phishing as an attack in which an attacker attempts to mislead an unsuspecting user and steal their money or personal information. He found that the victim is redirected to a fraudulent website mimicking a trusted site of the attacker through Email, SMS, or other digital means. Imgraben et al., (2014) proposed, that to avoid these attacks, users need to be trained to prevent phishing attacks from succeeding. However, Bulgurcu et al., (2010) found that in professional environments the training do not have the intended effect as the training feels as an obstruction. Gamification can be a solution here, by offering an engaging and interactive alternative, as proposed by research [1, 20, 27]. People in the age eighteen to twenty-five appear to fall for phishing scams the most often, as proposed by Sheng et al., (2010). This age group aligns most with the student population.

In this research, an interesting and engaging gamified approach will be presented to harden the Cyber Security Awareness of university students in phishing scenarios. A game for educational purposes, or a Serious Game as they are also called, was designed. The game was made specific to university students by evaluating which game elements are more interesting or engaging. This game was surveyed under university students to verify its effectiveness in preventing phishing success.

2. RELATED WORK

The literature studied in this paper was obtained from esteemed journals and conferences, such as ScienceDirect, Springer, ACM, and IEEE for the research. Related works were found using search terms such as 'gamification', 'security', 'cyber security', 'phishing', 'university students' and 'awareness'. Several documents were found which described research in this field. Below are the most relevant papers based on these search terms, summarized starting with 'gamification'. Continuing with specifying the age group and ending with this paper's proposal.

A systemic review of gaming technology for cyber security awareness reports that gamification can be an effective tool, as proposed by Alotaibi et al., (2016). The articles discussed in the review focused on general cyber security, phishing, and end-user safety. The paper selected twelve games that focused on cyber security awareness and training. These games showed promising positive results. However, these studies with games were only conducted in small sample groups and focused on children and teenagers.

Then Mayhorn et al., (2012) focused on educated people in the age range of 17-36. Proving that anti-phishing training was successful in lowering phishing susceptibility in the users.

On the psychological side of research, Baral et al., (2019) present a model to improve self-efficacy in phishing prevention. Maddux et al. proposed in 1995 that self-efficacy is the belief in the self that one can successfully complete a

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
36th Twente Student Conference on IT Febr. 4th, 2022, Enschede, The Netherlands.
Copyright 2022, University of Twente, Faculty of Electrical Engineering, Mathematics and Computer Science.

task. So when users believe they can successfully identify and cope with a phishing threat, the more likely they are to succeed.

Arachchilage et al., (2013 & 2014) described methods for phishing prevention focused on undergraduates from 18 to 25 years old, further specifying the age. Arachchilage's papers focused on the development of a model based on the Technology Threat Avoidance Theory proposed by Liang et al., (2009). This model proposes why IT users portrait avoidance behaviour. This model was then used to find a relation between self-efficacy and threat-avoidance behaviour by thwarting phishing attacks. Chiew et al., proposed in 2018 separate phishing categories such as spearfishing, smishing, and whaling.

The purpose of this paper is to design a methodology for designing gamified Cyber Security awareness training for preventing phishing against university students. However, as traditional training is ineffective as proposed by Bulgurcu et al., in 2010, there is a need for an interesting and engaging training to educate these users.

While there is sufficient knowledge out there on the function and application of gamification in phishing-specific Cyber Security awareness training, there are no games focusing on university students [1, 2, 3]. This disconnect is interesting as the age group 18-25 is already more likely to get phished compared to other age groups as proposed by Sheng et al., 2010. Existing training, such as the training provided by the University of Twente, forces repetition, which makes the training less enjoyable [17]. Does gamification have a positive effect on phishing prevention in university students when using an interesting and engaging gamified approach?

3. METHODOLOGY

This section presents the methodology for this research. The methodology has been specified for each sub-question, each subsection corresponds with a sub-question. See Figure 1 for the order of actions.

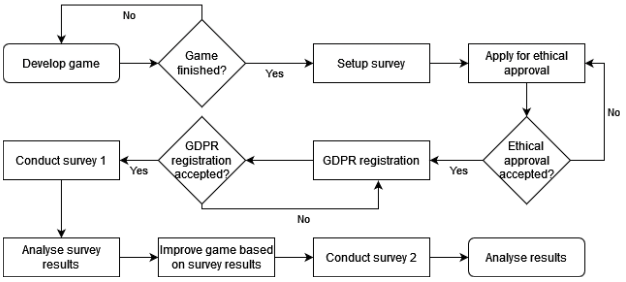


Figure 1: Flowchart of the Methodology

First, the focus has been on a literature study and games already out there. Based on the insights given by this literature, a game was created to evaluate its usefulness in students.

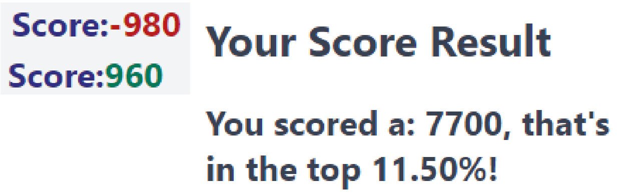
Second, a survey has been conducted among university students. University students are asked to play the game and subsequently fill in a survey. The data from this survey is analysed to improve the game.

Third, another survey using the game was conducted among university students. A subset of the university students have engaged with the gamified training. In the survey, the performance of students in various phishing scenarios were measured. For both surveys, ethical approval was obtained from the EEMCS ethical committee along with a GDPR registration [21, 19].

The goal of this research is to develop a game that is both engaging and interesting to university students, while not being less effective than regular phishing prevention training. The novelty of this research can be found in its form and competitiveness. The mobile-friendly web quiz format is not something that is used in gamified phishing training so far. Examples in research has opted for more of a game approach [24]. This game focuses on the competitiveness of users by providing them with aspects (i.e. score and time to compete with each other. Combined with the fact that each run will have a different order or different subset of questions allows for more replayability than other games out there.

3.1 Developing the game

In order to develop the game, common design patterns have been extracted from games and literature available [1, 2]. Patterns such as a time mechanic, a scoring mechanic with leaderboards and a way to provide positive/negative feedback while remaining educational.



Score:-980

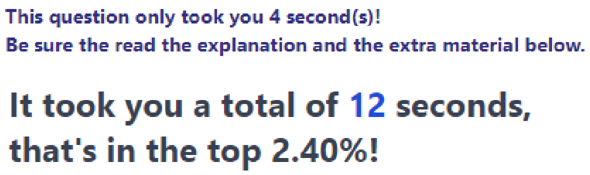
Score:960

Your Score Result

You scored a: 7700, that's in the top 11.50%!

Figure 2: Score mechanic screen-grabs

The score mechanic (see Figure 2) shows the user's current score throughout the game. It increases (or decreases) based on the amount of points, earned (or lost) per question. The score is displayed in green if it is above zero, or in red is it is below zero. At the end of the quiz, the total score is shown along with a (fictive) ranking.

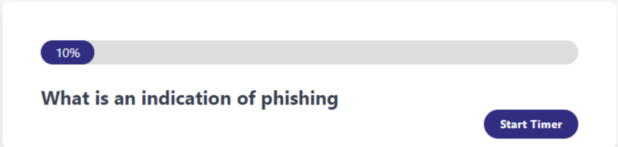


This question only took you 4 second(s)!
Be sure the read the explanation and the extra material below.

It took you a total of 12 seconds, that's in the top 2.40%!

Figure 3: Time mechanic screen-grabs

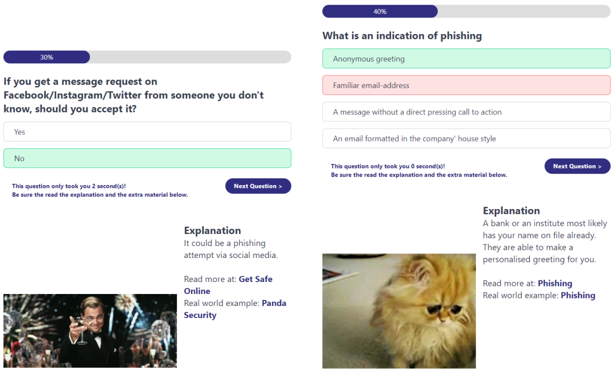
The time mechanic (see Figure 3) shows the time a user has to complete a question. At the time of design, a user is given a hundred seconds, this can be tweaked later. The timer is used to simulate a feeling of pressure in the user, similar to the feeling phishing tries to instil. The time taken to answer the question influences the score given. The faster a user answers, the higher the score. After each question, the user is shown how much time they took to answer the question. At the end screen, the user is shown how much time they took in total, along with a (fictive) ranking.



What is an indication of phishing

Figure 4: Progress mechanic screen-grab

The progress bar (see Figure 4) was implemented to give the users a sense of progress. In the initial design this bar was situated at the top of the page, however after feedback that a majority did not notice it was there. It has been included in the question view in order to be more visible.



If you get a message request on Facebook/Instagram/Twitter from someone you don't know, should you accept it?

Yes

No

This question only took you 2 second(s)!
Be sure the read the explanation and the extra material below.

Explanation
It could be a phishing attempt via social media.

Read more at: Get Safe Online
Real world example: Panda Security

What is an indication of phishing

Anonymous greeting

Familiar email-address

A message without a direct pressing call to action

An email formatted in the company' house style

This question only took you 0 second(s)!
Be sure the read the explanation and the extra material below.

Explanation
A bank or an institute most likely has your name on file already. They are able to make a personalised greeting for you.

Read more at: Phishing
Real world example: Phishing

(a) Positive feedback

(b) Negative feedback

Figure 5: Positive & negative feedback screen-grabs

The positive (Figure 5a) and negative (Figure 5b) feedback aspect are designed to educate a user while congratulating them when they have a question correct or in case of an incorrect answer to inspire them to do better. The correct answer of a question turns green, while the incorrect answer turns red. Additionally, a festive GIF is shown when the user. This GIF is a scene from the film The Great Gatsby, it displays the main character toasting while fireworks explode in the background. The negative feedback is a GIF of a kitten looking sad and bowing its head.

For the educational content over thirty questions were written along with explanations, additional sources, and where relevant with real world examples. If a user has a question wrong (see Figure 5b, they can look at the explanation why their answer was incorrect. If a user has a question correct (see Fi. 5a) but would like to learn more, they can follow the links to read additional material.

These patterns have been applied to designing a Serious Game to be used in the answering of the second and third sub-question. The game has been developed over the course of three weeks in four iterations. Each iteration increasing the gamification aspect, quality of the quiz, the visual feedback, or the educational value based on feedback from the supervisor.

3.2 Making it specific

In order to make the game specific to university students, a survey has been conducted on university students at the University of Twente. Before this survey could be conducted, ethical approval was acquired, and a GDPR registration was made in order to comply with the UT standards of research. An invitation for the survey was sent to several study WhatsApp groups, Discord servers & LinkedIn. Additionally, my supervisor & track chair assisted in reaching more university students. A group of twenty-nine students have been presented the game, the majority of students are from a computer science background. The students were subsequently asked to fill in a twenty-six questions survey. After a student granted consent to their data being used, the respondent could play the game. After playing the game, they filled in their score and time. This was followed by four questions asking what they (dis)liked about the game, with a text field for an explanation. This was structured as a multiple-choice question with the aspects of the game as options. These aspects are; the knowledge obtained in the game, the interactivity, the score mechanic, the time mechanic, the questions asked in the game, the design, the progress bar, the engagement of the game and finally an "other" option where the respondent could fill in a personal answer. Then, the survey asked to rank certain aspects of the game on a one to five scale [11], followed by a text field for an explanation. The survey proceeds by asking whether the score or the time motivated the participant to retake the quiz. Finally, the respondents are asked if they think this game would be useful in preventing phishing attacks in university students, with a field for an explanation. The results of this survey were used to improve the game.

3.3 Verifying effectiveness

In order to verify the game's effectiveness, a group of twenty-four university students has been split into two groups. One group, consisting of twelve students, played the gamified training developed in research question two. The survey started off by asking whether a student has already had a phishing training prior to the survey. It also asks how capable they perceive themselves in identifying phishing threats as well as their Dutch fluency, as 40% of the phishing attempts in the survey are in Dutch. The survey contained six phishing emails and SMS messages, as well as four legitimate emails and SMS messages. These messages have been collected from the researcher's email account and messaging app. The students have been asked to determine which is which. After each phishing or legitimate message, the participants are asked to explain their reasoning.

4. RESULTS

This section presents the results obtained from the conducted research. In the subsequent sections, the results of each of the research steps are discussed. The headers are the same headers as in the methodology section to indicate the link between the research question, its methodology and its results.

The first subsection describes the results of developing the game. It discusses already existing games and its differences. With the use of screenshots, it shows the result of the time invested in developing the game.

The second subsection analyses the survey conducted to improve the game. Participants were asked to rank several aspects of the game, as well as explaining what they like or dislike about the game. The results of this survey were used to improve the game.

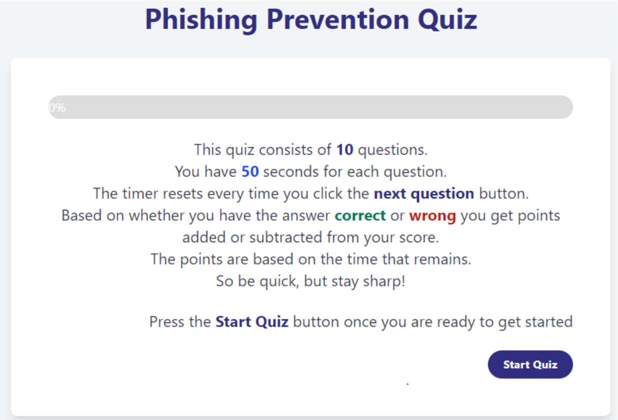
The third subsection analyses the survey conducted to evaluate the game's effectiveness. The participants were divided into two groups. On one side, the group who did the training, on the other side the group who did not. They were asked to judge their Dutch and phishing detection skills and then perform a phishing performance test.

4.1 Developing the game

Initially, the game developed by Google's Jigsaw team was considered [10]. Their approach is very interesting and engaging, however, their approach leans heavily on hovering over links. This does not work on mobile devices. An example from the other end of the spectrum, the FTC's phishing quiz has a very aged design and is not very interactive [8]. Bird's Life is the closest to this research in regard to target audience and way to reach them [24]. However, implementing a full 2D game did not fit this research' timeframe. While the game developed is not as involved as Bird's Life, it provides its value by achieving the interactivity by giving direct feedback each question and stay interesting by giving a new order of questions every attempt. While duplicates are unavoidable given enough retries, as there are thirty questions it will take several tries before all question are known to the user.

This game differs from training already out there, such as the University of Twente's (UT) own training, by using a score and time mechanic [17]. This mechanic encourages users to retake the quiz for a higher score or quicker time, compared to the UT training where the participant is forced to retake the quiz when he does not score above 80%. The score and time mechanic combined with its ranking system was designed to inspire competitiveness. Upon retaking the quiz for a higher score or faster time, the user would encounter a different set of questions. By repetition, the user is exposed to more knowledge and thus will learn more while going for the high score. The game developed is a mix between the interactivity of Bird's life [24] and educational value of Google's training [10] together with the low barrier of the FTC' training [8].

The game was developed in JavaScript using the Vue library [26]. To make the quiz specific to students in the Netherlands, several questions include institutions such as DUO, the IND and the University of Twente. The website was hosted at a server of the UT.



Phishing Prevention Quiz

0%

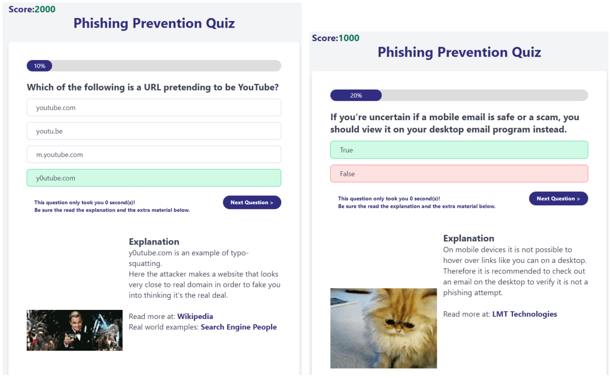
This quiz consists of **10** questions.
You have **50** seconds for each question.
The timer resets every time you click the **next question** button.
Based on whether you have the answer **correct** or **wrong** you get points added or subtracted from your score.
The points are based on the time that remains.
So be quick, but stay sharp!

Press the **Start Quiz** button once you are ready to get started

Figure 6: Introduction screen of the gamified phishing training

The game starts with a screen explaining the game, see Figure 6. The game has thirty questions, of which ten are randomly selected every round. Every question has two to four answers, with one correct answer. For each question, the participant has a hundred seconds to answer, in order to provide the participant with ample time to read the question and its answers.

If the user selects the correct answer, they receive a score equal to the amount of seconds left multiplied by ten. If incorrect, this amount is subtracted from his score. After each question, the user is presented with an explanation of why the answer was the correct one, along with some literature where they can read more. Depending on the question, this explanation also includes a real-world example.



Score:2000

Phishing Prevention Quiz

10%

Which of the following is a URL pretending to be YouTube?

youtube.com

youtu.be

m.youtube.com

y0utube.com

This question only took you 0 second(s)!
Be sure the read the explanation and the extra material below.

Explanation
y0utube.com is an example of typo-squatting.
Here the attacker makes a website that looks very close to real domain in order to fake you into thinking it's the real deal.

Read more at: **Wikipedia**
Real world examples: **Search Engine People**

(a) Screen of a correct answer in the gamified phishing training

Score:1000

Phishing Prevention Quiz

20%

If you're uncertain if a mobile email is safe or a scam, you should view it on your desktop email program instead.

True

False

This question only took you 0 second(s)!
Be sure the read the explanation and the extra material below.

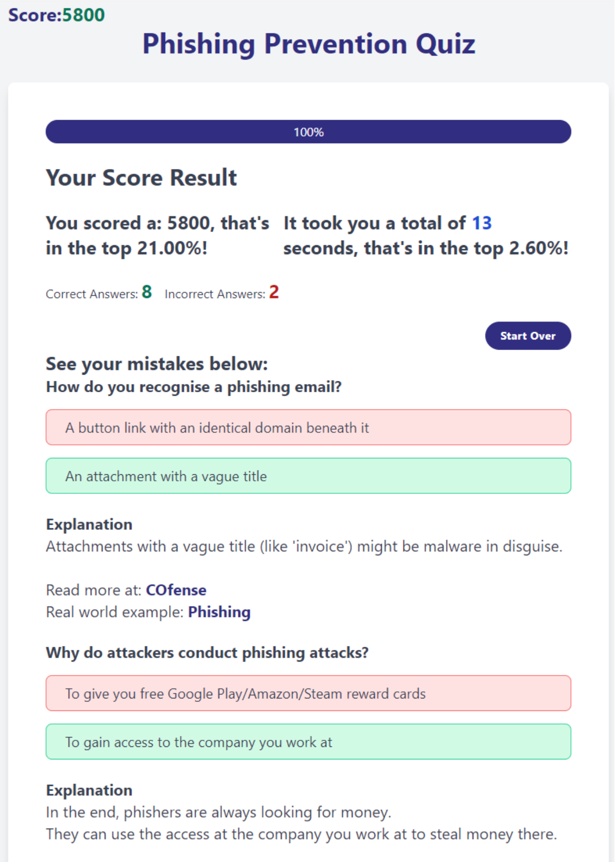
Explanation
On mobile devices it is not possible to hover over links like you can on a desktop. Therefore it is recommended to check out an email on the desktop to verify it is not a phishing attempt.

Read more at: **LMT Technologies**

(b) Screen of an incorrect answer in the gamified phishing training

Figure 7: Screenshots from a correct and incorrect answer

Additionally, a fireworks or sad kitten GIF is shown to the user based on whether they answered correctly, see figures 7a and 7b respectively.



Score:5800

Phishing Prevention Quiz

100%

Your Score Result

You scored a: 5800, that's in the top 21.00%!

It took you a total of 13 seconds, that's in the top 2.60%!

Correct Answers: 8 Incorrect Answers: 2

See your mistakes below:

How do you recognise a phishing email?

A button link with an identical domain beneath it

An attachment with a vague title

Explanation
Attachments with a vague title (like 'invoice') might be malware in disguise.

Read more at: **COfense**
Real world example: **Phishing**

Why do attackers conduct phishing attacks?

To give you free Google Play/Amazon/Steam reward cards

To gain access to the company you work at

Explanation
In the end, phishers are always looking for money.
They can use the access at the company you work at to steal money there.

Figure 8: Final screen of the gamified phishing training

After answering the ten questions, the user is shown a summary screen. This screen shows their final score and total time, as well as a fictive ranking. Additionally, all their mistakes are collected here, so the user can review them. This screen can be seen in Figure 8.

The framework for the game has been set up in such a way it should be trivial to add or rewrite questions. This way, future research could build off the groundwork that has been laid here.

4.2 Making it specific

In the survey the participants scored an average of 6300 points, as shown in Figure 9, in 93 seconds, as can be seen in Figure 10.

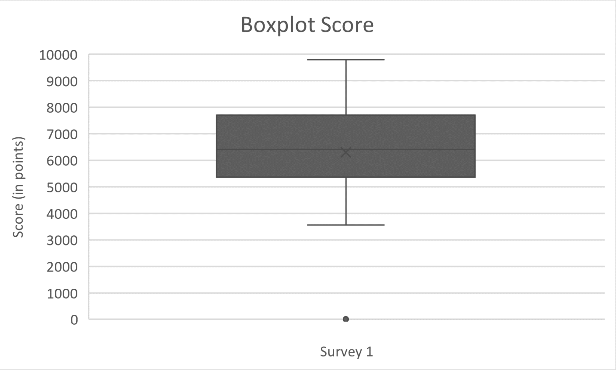


Figure 9: Survey participants' scores

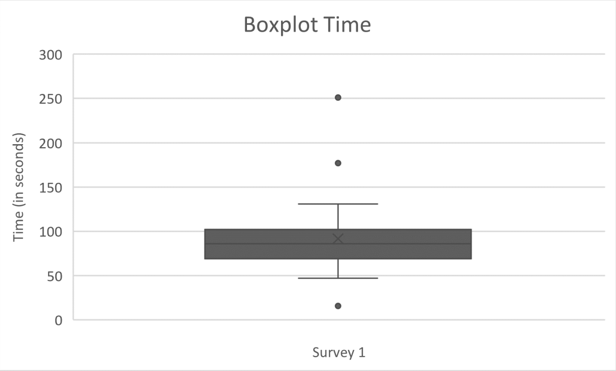


Figure 10: Survey participants' times

Figure 11: Statistic of the participants responding to survey 1

The y-axis shows the score in points or the time in seconds, respectively [25]. As can be seen in the score graph, the participants score close to the theoretical maximum. A persons' score lies between positive and negative ten thousand points. This may be due to the questions being too easy, something that was suggested by several participants. It is also very likely that the participants are already familiar with phishing material, as they are from the Computer Science background. In the time graph, we can see that the hundred seconds limit per question was superfluous. Every participant took less than thirty seconds per question. Following this observation, the time allotted for each question was lowered to fifty seconds. This amount was selected as it is still above thirty, as several participants mentioned they felt the time pressure.

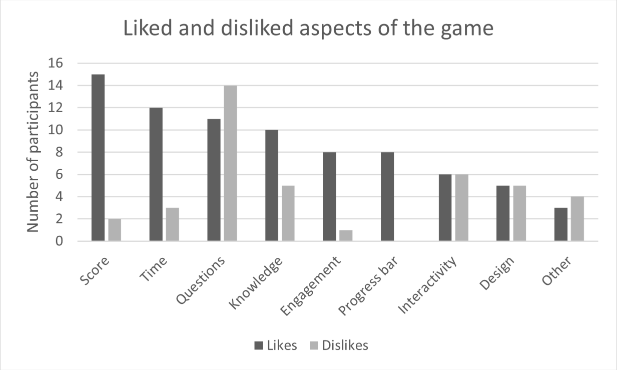


Figure 12: Likes and dislikes of the game's aspects

The scoring mechanic was the most liked aspect by 52% of the participants, as can be seen in Figure 12, followed by the time mechanic (48%) and the questions asked (38%). The most disliked feature was the questions asked by 48%, followed by the game's interactivity (21%) and knowledge obtained through the game (17%). The low score of the questions can be attributed to their difficulty, as 38% of the participants found the questions too easy. This is most likely because they are Computer Science students themselves, as suggested by 36% of this subset, and thus are already familiar with most of the material presented in the game.

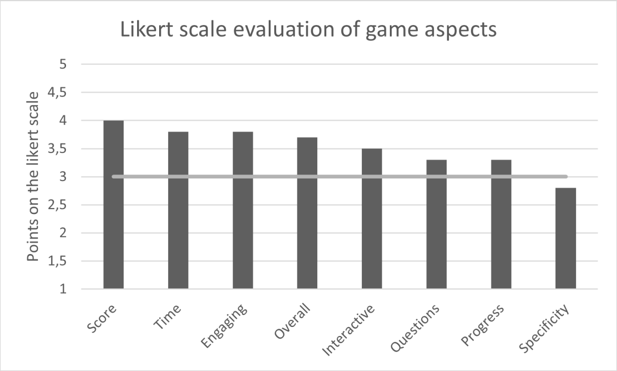


Figure 13: Averages of the Likert scales

The y-axis displays the average score on the Likert scale, with one being the lowest and five being the highest. The x-axis shows the aspects of the game, sorted from highest to lowest score. The line in the graph at 3 shows the middle point of the Likert scale. In Figure 13 the overall score of the game was a 3.7, indicating that the game is good but has room for improvement. As can be seen in the figure, all aspects of the game performed above average, except for specificity. This aspect is surveyed as to how targeted the participants felt as university students by the game. It follows that the questions used in this game should be made more specific to university students. From the explanation given for all the likes, dislikes, and Likert scores, several suggestions were extracted, see Table 1.

These suggestions were ranked by the amount of times they were suggested by different participants. Some suggestions included using different correct or wrong GIFs, making questions more specific to university students, implementing different question types such as multiple-choice, sliders, or image questions. Development time was focused on reviewing existing questions, improving explanations and their visibility.

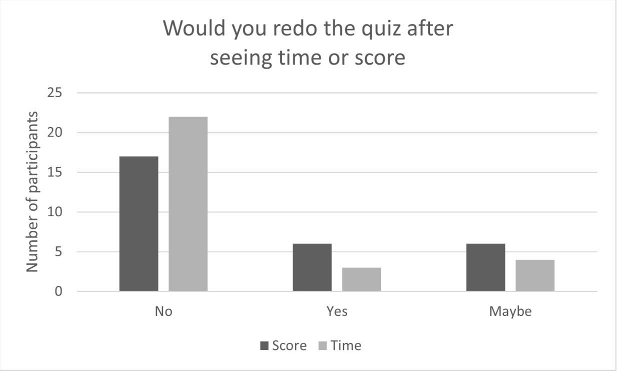


Figure 14: Participants interested in redoing the quiz

When looking at the replayability of the quiz, most respondents stated that the questions were too easy to motivate them to retake the quiz. This is most likely because most participants are from a Computer Science background and are thus already familiar with phishing material. As can be seen in Figure 14, the score was a more motivating reason to redo the quiz than the time.

4.3 Verifying effectiveness

The second survey was filled in by twenty-four students. The students have been divided into two groups of twelve to gauge the effectiveness of the gamified phishing training. One group played the gamified phishing training, the other group did not.

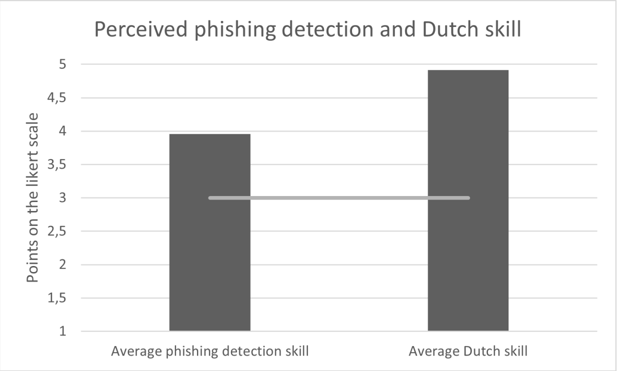


Figure 15: Perceived ability

When asked to grade their perceived phishing ability and Dutch skills on a one to five scale [11], everyone responded above average. The results can be seen in Figure 15. Of the twelve students that did not play the training, four students (33%) followed some kind of phishing training before. 50% of them followed the UT training, the remaining two followed a training for work or a generic training online.

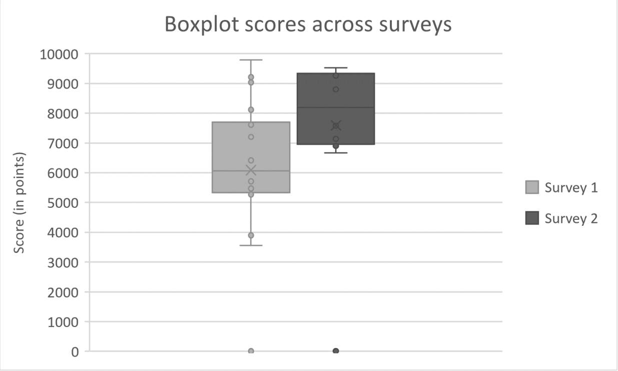


Figure 16: Survey participants' scores for survey 1 & 2

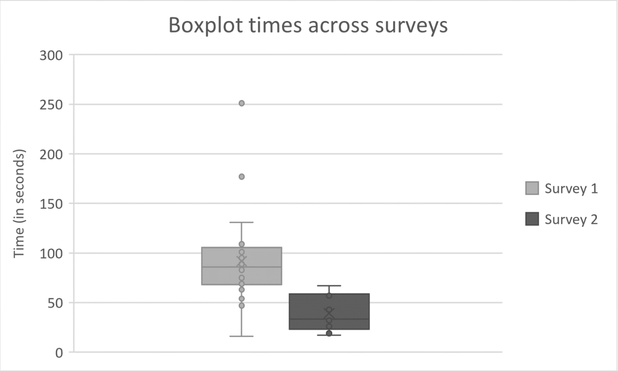


Figure 17: Survey participants' times for survey 1 & 2

Comparing the results of the phishing game scores and times, shows an improvement among participants. The average score was 7595 points, an increase of about 1300 point (or 21%) compared to the previous scores as can be seen in Figure 16. The average time was 39 seconds, a decrease of 52 seconds (or 57%) compared to the previous survey in Figure 17. The steep decrease in time can be explained by the redesign of the time feature. The time taken for reading the question is no longer counted for the total time, as requested by a participant with Dyslexia. Based on this average, the time per question can be reduced even further, as it was only reduced to fifty between the surveys.

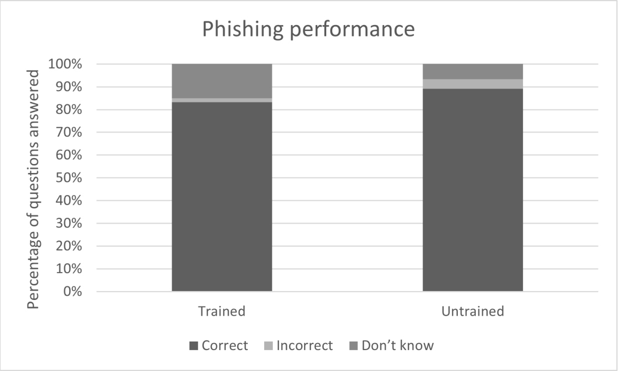


Figure 18: Phishing performance per group

All of them were presented with the same ten phishing questions. The performance difference between the trained and the untrained group lie within the margin of error, as can be seen in Figure 18. This slight difference can be explained by the fact that most of the survey participants have been sourced from students with a computer science background. These students are already familiar with phishing material and benefit little from a phishing training. While the question answered correct is slightly higher in the untrained group, their questions answered incorrect is also slightly higher. As a consequence, the error rate of the untrained group is 4.2%, opposed to 1.7% of the trained group [5]. The full contingency table and its classification report can be found in Tables 2 & 3 respectively. In conclusion, these results are so close that it can be concluded that this training need to be improved further in order to make a difference.

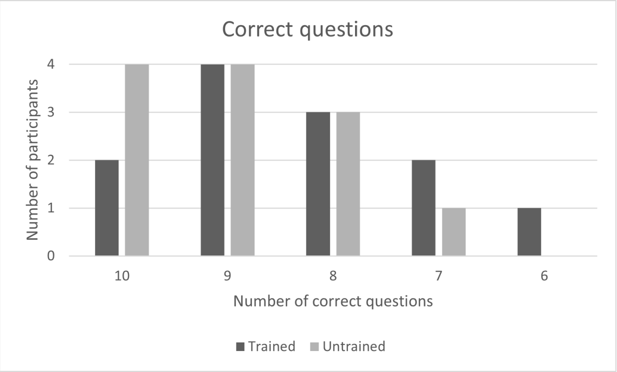


Figure 19: Number of questions correct by training

Overall, the performance of both groups is very close when looking at Figure 19. This graph shows how many participants had the same number of correct question separated by the trained and untrained group. The untrained group had two members more who performed flawlessly. The trained group had one member more who had six & seven questions correct. In the other categories, the trained and untrained group performed identical.

5. DISCUSSION

This research focused on university students in the Netherlands, as university students are the most likely victims in their age range [22]. This research also restricted itself to phishing. The game has been designed in JavaScript using NodeJS. The website to play this game has been designed Mobile-first, as to cater to mobile and tablet users as well. Supporting these devices has been limited to Android and iOS. This game is supported in Chrome (and its chromium derivatives) and Firefox.

Initially, this research focused on international students. However, not enough students responded to make the study statistically significant. Hence, the research was broadened to all university students. This created a problem for the novelty of the research, as gamified phishing training has been conducted among university students before [24]. This research is still unique, as the game developed is more of a web quiz than a game, focusing on replayability and competitiveness.

Little research has been done on the prevention of phishing for university students. Most research in this field focuses on children and teenagers. Due to the time constraints of this research, the phishing training is a proof of concept and focuses on university students in the Netherlands. Obtaining ethical approval was another delay that was encountered. The back and forth with questions and clearing up uncertainties unfortunately delayed obtaining approval significantly. As a result, the survey started later than planned.

Development of the platform took more time than anticipated, as this project was the researcher's first time doing JavaScript and front-end design. While JavaScript's learning curve is relatively low, combining it with a new library that he had not worked with before as well as the missing insider knowledge of CSS made the development slower than initially anticipated. This also influenced the choice of feedback implemented extracted from the first survey. In case the researcher was more comfortable with JavaScript, he could have implemented more of the feedback.

Following the suggestions extracted from the first survey, the next iteration of the game should consider a difficulty system. This difficulty system could ask a question in the same knowledge area when a participant answers a question wrong. Additionally, the difficulty system could prioritize asking more difficult questions if the participant has multiple correct questions in a row and vice versa. The next version could also support different question types; multiple-choice, image questions, sliders. Implementing these question types can increase the interactivity. Additionally, the game developed during this research could be improved further by making the social features more robust. By adding leaderboards, high scores, or a way for participants to challenge their friends via email or social media.

In the second survey, only one respondent reported their Dutch skill below a 5. In order to draw conclusions on the relation between Dutch skill and phishing performance, the survey should include more non-Dutch respondents. The second survey shows the need for a larger sample size, especially a sample set outside the computer science space. Familiarity with phishing material most likely influences phishing performance. However, as no data on educational background was collected, no conclusion could be drawn. Future work should also consider measuring phishing performance by gender, so the game can be made even more specific.

6. CONCLUSION AND FUTURE WORK

The literature contains an abundant body of research into the application of gamification for Cyber Security Awareness and phishing prevention. However, a study focusing on an interesting and engaging gamified approach on phishing against university students has not yet been conducted using a web quiz format. The framework developed can be used for future research. The game that has been developed is well liked while having room for improvement. It is a more interactive and engaging alternative to regular phishing training.

Future work should focus on expanding the game with features such as different question types and a difficulty scaling system. These features will enhance the interactivity and engagement further. As whether the game contributes to improved phishing performance is still unclear, future work should focus on conducting a larger scale survey with students outside the computer science focused background.

Acknowledgments

I would like to thank D. K. Sarmah for her guidance as supervisor during this research. This research would not have gone as smoothly without the help from Network Services and Network Security track chair S. Bayhan. I would also like to thank B. Meijerink for providing hosting for the gamified training website. R. Bezem deserves a special thanks for his LaTeX assistance. As well as B. van Vliet & T. Keijser for assisting with my Javascript/Vue questions. Additionally, I would like to extend my gratitude to L. Kreijne, L. & J. Schöls and R. Langbroek for alpha-testing my game. A special thanks to R. de Ridder & G.H. Nijland for proofreading my paper. Finally, I would like to extend thanks to everyone who contributed by filling in the survey.

References

[1] Faisal Alotaibi et al. "A review of using gaming technology for cyber-security awareness". In: *Int. J. Inf. Secur. Res.(IJISR)* 6.2 (2016), pp. 660–666.

[2] Nalin Asanka Gamagedara Arachchilage and Steve Love. "A game design framework for avoiding phishing attacks". In: *Computers in Human Behavior* 29.3 (2013), pp. 706–714.

[3] Nalin Asanka Gamagedara Arachchilage and Steve Love. "Security awareness of computer users: A phishing threat avoidance perspective". In: *Computers in Human Behavior* 38 (2014), pp. 304–312.

[4] Gitanjali Baral and Nalin Asanka Gamagedara Arachchilage. "Building Confidence not to be Phished Through a Gamified Approach: Conceptualising User's Self-Efficacy in Phishing Threat Avoidance Behaviour". In: *2019 cybersecurity and cyberforensics conference (CCC)*. IEEE. 2019, pp. 102–110.

[5] Michael W Browne. "Cross-validation methods". In: *Journal of mathematical psychology* 44.1 (2000), pp. 108–132.

[6] Burcu Bulgurcu, Hasan Cavusoglu, and Izak Benbasat. "Information security policy compliance: an empirical study of rationality-based beliefs and information security awareness". In: *MIS quarterly* 31.3 (2010), pp. 523–548.

[7] Kang Leng Chiew, Kelvin Sheng Chek Yong, and Choon Lin Tan. "A survey of phishing attacks: Their types, vectors and technical approaches". In: *Expert Systems with Applications* 106 (2018), pp. 1–20.

[8] FTC. *FTC Phishing Quiz*. `https://www.ftc.gov/tips-advice/business-center/small-businesses/cybersecurity/quiz/phishing`. Last accessed on 2022-01-20.

[9] James Imgraben, Alewyn Engelbrecht, and Kim-Kwang Raymond Choo. "Always connected, but are smart mobile users getting more security savvy? A survey of smart mobile device users". In: *Behaviour & Information Technology* 33.12 (2014), pp. 1347–1360.

[10] Jigsaw. *Phishing Quiz by Google, Jigsaw*. `https://phishingquiz.withgoogle.com/`. Last accessed on 2022-01-20.

[11] Ankur Joshi et al. "Likert scale: Explored and explained". In: *British Journal of Applied Science & Technology* 7.4 (2015), p. 396.

[12] James Lance. "Chapter 1 - Banking on phishing". In: *Phishing Exposed*. Syngress, 2006, pp. 1–35. ISBN: 978-1-59749-030-6. DOI: `https://doi.org/10.1016/B978-159749030-6/50006-4`.

[13] Huigang Liang and Yajiong Xue. "Avoidance of information technology threats: A theoretical perspective". In: *MIS quarterly* (2009), pp. 71–90.

[14] James E. Maddux. "Self-Efficacy Theory". In: *Self-Efficacy, Adaptation, and Adjustment: Theory, Research, and Application*. Boston, MA: Springer US, 1995, pp. 3–33. ISBN: 978-1-4419-6868-5. DOI: `10.1007/978-1-4419-6868-5_1`. URL: `https://doi.org/10.1007/978-1-4419-6868-5_1`.

[15] Christopher B Mayhorn and Patrick G Nyeste. "Training users to counteract phishing". In: *Work* 41.Supplement 1 (2012), pp. 3549–3552.

[16] Susan Morrow and Nick Maynard. *Online Payment Fraud: Emerging Threats, Segment Analysis & Market Forecasts 2021-2025*. `https://www.juniperresearch.com/researchstore/key-vertical-markets/online-payment-fraud-research-report`. Last accessed on 2021-11-25.

[17] Wim Olijslager. *UT Security Training*. `https://www.utwente.nl/en/cyber-safety/news/2021/10/35194/take-the-cyber-security-course-and-learn-whether-your-data-is-protected-properly`. Last accessed on 2022-01-17.

[18] Swapan Purkait. "Phishing counter measures and their effectiveness–literature review". In: *Information Management & Computer Security* 20.5 (2012), pp. 382–420.

[19] UT Research GDPR Registration. *UT Research GDPR Registration*. `https://www.utwente.nl/en/bms/datalab/research-data-and-gdpr/research-registration/`. Last accessed on 2022-01-26.

[20] Michael Sailer et al. "How gamification motivates: An experimental study of the effects of specific game design elements on psychological need satisfaction". In: *Computers in Human Behavior* 69 (2017), pp. 371–380.

[21] Ethics Committee Computer & Information Science. *Ethics Committee Computer & Information Science*. `https://www.utwente.nl/en/eemcs/research/ethics/`. Last accessed on 2022-01-26.

[22] Steve Sheng et al. "Who Falls for Phish? A Demographic Analysis of Phishing Susceptibility and Effectiveness of Interventions". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '10. Association for Computing Machinery, 2010, pp. 373–382.

[23] Camille Singleton. *X-Force Threat Intelligence Index 2021*. `https://www.ibm.com/downloads/cas/M1X3B7QG`. Last accessed on 2021-11-25.

[24] Patrickson Weanquoi, Jaris Johnson, and Jinghua Zhang. "Using a game to improve phishing awareness". In: *Journal of Cybersecurity Education, Research and Practice* 2018.2 (2018), p. 2.

[25] David F Williamson, Robert A Parker, and Juliette S Kendrick. "The box plot: a simple visual method to interpret data". In: *Annals of internal medicine* 110.11 (1989), pp. 916–921.

[26] Evan You. *VueJS*. `https://vuejs.org/`. Last accessed on 2022-01-21.

[27] S V Zenkina et al. "Capabilities of digital gamification resources to form the basis of information security". In: *Journal of Physics: Conference Series* 1691 (Nov. 2020), p. 012064. DOI: `10.1088/1742-6596/1691/1/012064`. URL: `https://doi.org/10.1088/1742-6596/1691/1/012064`.

APPENDIX

Survey 1 Suggestions

Suggestions ranked by votes		
Suggestion	Number of votes	Implemented
Make the questions more specific	18	✗
Make the progress bar more visible	13	✓
Add different types of questions	13	✗
Improve explanations & their visibility	12	✓
Increase question difficulty	11	✗
Change the design	7	✗
Reduce time pressure	7	✓
Improve quiz flow	7	✗
Review existing questions	6	✓
Reduce time per question	4	✓
Implement leaderboards	2	✗

Table 1: Suggested improvements extracted from Survey 1

Survey 2 Tables

			Predicted		
			Phishing	Legimate	Total
Observed	Trained	Phishing	60	0	60
		Legitimate	2	40	42
		Don't know	10	8	18
	Untrained	Phishing	63	2	65
		Legitimate	3	44	47
		Don't know	6	2	8
		Total	144	96	240

Table 2: Contingency table of phishing performance in Survey 2

	Error rate	Accuracy	Sensitivity	Specificity	Precision	False Positive Rate
Trained	0,017	0,833	1	0,952	0,968	0,048
Untrained	0,042	0,892	0,969	0,936	0,955	0,064

Table 3: Classification report of phishing performance in Survey 2