Password: A Serious Game to Promote Password Awareness and Diversity in an Enterprise

Gokul Chettoor Jayakrishnan, Gangadhara Reddy Sirigireddy, Sukanya Vaddepalli, Vijayanand Banahatti, and Sachin Premsukh Lodha, *TCS Research, Tata Consultancy Services Limited, Pune, India;* Sankalp Suneel Pandit, *Former employee of TCS Research, Tata Consultancy Services Limited, Pune, India*

https://www.usenix.org/conference/soups2020/presentation/jayakrishnan

This paper is included in the Proceedings of the Sixteenth Symposium on Usable Privacy and Security.

August 10–11, 2020

978-1-939133-16-8

Open access to the Proceedings of the Sixteenth Symposium on Usable Privacy and Security is sponsored by USENIX.

Passworld: A Serious Game to Promote Password Awareness and Diversity in an Enterprise

Gokul Chettoor Jayakrishnan, Gangadhara Reddy Sirigireddy, Sukanya Vaddepalli, Vijayanand Banahatti, Sachin Premsukh Lodha, Sankalp Suneel Pandit[1]
TCS Research, Tata Consultancy Services Limited, Pune, India ([1]former employee)
{gokul.cj, gangadhara.sirigireddy, sukanya.vaddepalli, vijayanand.banahatti, sachin.lodha}@tcs.com, pandit.sankalp123@gmail.com

Abstract

Usage of weak passwords for authentication within an organization can be exploited during cyberattacks leading to unauthorized account access, denial of service, data and identity theft, sabotage etc. Such attacks could bring financial and reputational losses apart from legal consequences. Organizational password policies came into being in an attempt to encourage users to create more complex and diverse passwords. However, it has been observed that people show similar behavior in adopting those policies and end up creating passwords with similar patterns. Security training has been found to be a popular mechanism in an enterprise setting, of which, game-based trainings have shown positive impact with an added advantage of being immersive. In this paper, we present a serious game-based training on creating password security awareness among enterprise users. The training involves promoting understanding among users about various common password heuristics during password creation. This study focuses on two research questions: 1) Can a game-based password awareness training teach participants about the various password heuristics? 2) Can such a training improve the organizational password diversity? With a participation of 4,906 employees from our enterprise in the study, we were able to observe effects of game-based training on password awareness. We also found insights during the study to show that users created diverse passwords.

1. Introduction

Despite advancements in user authentication methods, the decades old system of username-password combination is still prevalent [54, 57]. Even the method of two-factor authentication generally consists of passwords as one of its factors [3, 40]. The human element involved in password creation is one of the major factors affecting password strength [35]. Studies show that people are more likely to use weaker and easily memorable passwords because of the lack of knowledge in creating stronger passwords [20] or due to the limitations in memorizing passwords [56]. This makes it easier for attackers to crack the passwords, resulting in security breaches and loss of personal and confidential information. Organizations have had their fair share of difficulties in dealing with password breaches [12, 30, 45, 46]. In an enterprise scenario, attackers often target individuals to steal passwords and gain access to organizational data [6]. Verizon's Data Breach Investigation Report (DBIR) [52] states that 63% of data breaches worldwide happened due to weak or stolen passwords or usage of default passwords. Organizational password policies such as setting a password expiration period could also lead to adoption of insecure methods like writing down the passwords [19] or using newer passwords with minimal difference from the old ones. For example, appending symbols and/or digits at the end [37] could result in weak and similar passwords. These fall under insecure password practices and avoiding them will help create better passwords.

Password meters are found to be useful in creating stronger passwords [51]. However, most of them provide basic feedback [13, 51] and many of them also rate passwords inconsistently [39]. There is a need to create awareness among users when it comes to creating better passwords. An advanced data-driven password meter by Ur, et al. [48] that uses several heuristics to score passwords showed compelling results, but with limitations.

Among various training methods, text-based training like reading documents has been said to be monotonous [10]. However, interactive games are found to be helpful and immersive [8, 41, 53] when it comes to training in cybersecurity. Considering this, we intended to provide an interactive training experience to the employees of our

USENIX Symposium on Usable Privacy and Security (SOUPS) 2020. August 9 -- 11, 2020, Virtual Conference.

organization by following a game-based password awareness methodology. We utilized the heuristics from the previous study [48] to create a comprehensive training method in order to teach password practices. Heuristics, in the context of this study, denote the techniques or practices that have to be satisfied to improve the overall password strength. For example, a heuristic like "Not more than three consecutive repeating characters" will be satisfied if the created password does not have more than three consecutive repeating characters. We conducted the game-based experiment as a part of our organization's annual information security awareness week. The online 2D web game titled Passworld focuses on enterprise password training, with little compromise on the fun element of gameplay. This paper details the experiment, results, and observations obtained from our study.

2. Related Work

There are several previous studies relating to user behaviors, tools for password security, and game-based cybersecurity awareness training.

2.1. Password Security: User Behaviors

When it comes to user behavior, studies [2, 50] show that users assess various cybersecurity threats by themselves and their assessment often ends up being inaccurate, mainly because they are not knowledgeable in this regard. The usage of multiple online accounts leads to password reuse [2]. This, combined with insecure creation practices, make passwords all the more vulnerable. The users' cognitive ability to simultaneously deal with multiple passwords is limited [23]. A recent study on users' perceptions on password strength [50] shows that many users' idea of strong passwords, like adding digits and symbols at certain positions, is exploited by today's password cracking tools. Another study [49] shows that even with the awareness about importance of creating strong passwords, many users create passwords that are easy to guess. While many users did not know how to create strong passwords, many believed that they were creating strong passwords, but in fact, those passwords were just predictable [49]. As per Habib et al. [20], in workplaces with a password expiration policy, many users tend to use password creation strategies that are very predictable and with minimal modifications. A study with adaptive password blacklisting policies [38] was found to have good results. However, the study also advises using a good password-composition policy to augment the structure-based adaptive, blacklist system. Certain tools exist to help users create better passwords.

2.2. Existing Tools for Password Security

Proactive password checking is a method of checking a user's password to see if it satisfies certain criteria to consider it as secure [23, 56]. Password strength meters implement this by checking several criteria for testing the strength of a password. A recent study with a data driven password meter [48] showed positive results, which were attributed to the feedback mechanism implemented into the tool. However, the study also concluded that the overall effect of the meter was not strong and some participants did not trust suggestions from a computer. Studies on users' perceptions on password systems and behaviors [2, 55, 57] have recommended the need to provide training and instructions on constructing usable and secure passwords, and to provide adequate feedback during training to enhance their knowledge.

2.3. Existing Password and Security Awareness Trainings

A report by security firm LastPass in 2018 [24] revealed that only 19% of users create strong passwords at work and 62% of users use the same passwords for work and personal accounts. It implies that knowing an employee's personal password makes it easier for hackers to guess their work password [24]. This suggests the importance of making enterprise users aware of the various password heuristics that could help them create secure passwords for both personal and work accounts. Many enterprise-training methods like educational documents, mail-based embedded training [29] etc. are available. However, these are considered passive training methods that are found to be less engaging and monotonous [10]. On the contrary, interactive games facilitate participation between the teaching agent and learner [32, 53].

2.4. Interactive Training: Serious Games

Games with a purpose, titled serious games, perform an additional task of creating awareness apart from entertainment [5]. Serious games have been used to spread cybersecurity awareness with successful results. Previous studies with games like Anti Phishing Phil [41], CyberCIEGE [44], Phishy[8], GAP [47], Cyberaware [18], PASDJO[39], and Control-Alt-Hack [14] have shown that games are not only effective in training cybersecurity concepts, but are also more engaging, compared to the traditional methods like reading documents.

Considering the existing password awareness games, PASDJO [39] trains users on measuring the strengths of on-screen passwords through the method of inspection. GAP [47] focuses on training users on predictable positions of uppercase letters, digits, and symbols. While these game-based studies provided good results in terms of user's understanding when it came to insecure password

practices, they focused on a very small set of password heuristics. We tried to extend this study to a larger audience. We followed a game-based approach for password training that focused on a set of 16 password heuristics. To bring the game more in line with the latest findings and studies, we used the results from the study by Ur, et al. [48] and a set of password recommendations from a study by Vu, et al. on improving password security and memorability [54]. These were analyzed in light of the organizational password complexity requirements, and were organized to form a set of heuristics to be taught through the game. The tool used in a previous study [38], suggests modifications to passwords entered by users in order to make them compliant with policies. Our intent was to test the user understanding by letting them apply their learnings about the heuristics, to their passwords without any additional suggestions on modifications to the entered password. The reports from our organization's Corporate Security Team (CST), which takes care of awareness training, suggested from their data that employees participate in mandatory training modules because of the condition that they must finish it. This condition could provide a sense of urgency to complete, rather than understand the course content with a curiosity to learn. So, as suggested by CST, we decided on a non-mandatory, gamified way of training to create awareness about passwords among participants. The training was conducted in collaboration with CST, which also took care of all the clearances and ethics review. The game was made part of the normal workday.

3. Passworld Game

We developed an online, web-based game to spread awareness about various password heuristics among the employees of our organization. Our unique game titled "Passworld" follows certain basic principles to ensure that an optimal learning outcome is obtained. We followed a set of established theories and principles during the design phase of our game.

Experiential Gaming Model: Games have been shown to be most successful and engaging when they provide a flow experience [11] to players [26]. An optimal flow experience, combined with experiential learning and feedback, termed as the Experiential Gaming model, is found to maximize the impact of a training game [26]. Passworld tries to incorporate a similar learn and reflect methodology with emphasis on immersion and enjoyment by providing learning tips, timely feedback, and an easy-to-learn interface, similar to the ones in the classic and widely famous games of the 80s and 90s like Super Mario Bros. [34] and Adventure Island [21].

Bloom's Taxonomy: The game also follows all the six levels of Bloom's Taxonomy in the cognitive domain [27, 42] to maximize the learning outcome. The game instructions, learning tips, and basic information on password heuristics form the first level *Knowledge* in the Bloom's hierarchy. The game presents these learning materials in a fun and entertaining way to help with *Comprehension*. Players have to apply this knowledge while creating passwords, which satisfies *Application* level in the hierarchy. Further, players *Analyze* each password heuristic and *Synthesize* their own knowledge by adding up pieces of information obtained from the game levels. The instant and delayed feedback in the game helps them to *Evaluate* their choices, thereby helping in reflection, a learning science principle [15].

Password also follows the **Conceptual-Procedural** gaming principle [41] to ensure a better learning outcome. This principle states that conceptual and procedural knowledge augments one another in an iterative process [25]. This, along with game design patterns like integration pattern, cognition pattern, and presentation pattern [5], helped us decide a flow and overall design for the game. We also ensured that the Six "I" Framework of Serious Game Design [4] that focuses on Identity, Immersion, Interactivity, Increased Complexity, Informed Teaching, and Instructional Content, is followed.

3.1. Game Description

Passworld is a 2D, single-player, horizontal scrolling platformer game [9] falling in the Action-Adventure [33] genre, developed in the Unity3D game engine. We tried to keep the game as lightweight as possible for compatibility with all the browser versions and machines.

3.1.1. Game Design Choices

We decided on providing a positive gameplay experience by intertwining password awareness with a platformer-based game, where both intrinsic motivation of fun element of gameplay as well as extrinsic motivation like rewards and benefits merge [17]. Passworld was made to have a gameplay experience similar to that of certain classic games [21, 34] to create a positive mindset in players, through a feeling of nostalgia [7]. We chose a horizontal, jungle-based platformer game considering the non-monotonous gameplay factor, game time, interactions and visuals, and the analogy that reflects real-life scenario of password attacks. We tried to relate the real world to our game through the following design choices:

Open-interconnected world: The real cyber world is always open and interconnected. The chances of cyber-attacks and password breaches are high, if we do not follow proper security measures. We wanted to create a similar scenario within the game. Therefore, we chose a jungle

environment, which is a wide-open scenario. If the user is not careful, the chances of being attacked by animals are much higher.

Digital assets need protection where passwords play an important role. Similarly, the game focuses on storing important ancient artifacts using secure gates.

Know thy enemy: In order to create secure passwords, one must know the weaknesses of passwords that are exploited by attackers. This is where the game introduces various animals. The animals in the game provide tips about several password heuristics. They also check the user created passwords to see if heuristics are satisfied, and if not, they attack. We used the password heuristics from a previous study [48]. While [48] was based on using the heuristics to check and analyze user created passwords, and provide suggestions for improvement, we used the heuristics to teach users about individual password requirements.

Prepare to defend: In real life, we can create passwords using all available character classes. We converted this into a game resource, where players have to gather the different character classes to create their passwords (gates). The resources are the raw materials used to create the gates, just like L, U, D, and S character classes are used to create a password. Typical enterprise password policies mandate usage of all character classes while creating passwords.

Build a strong defense: Passworld teaches users about password heuristics and requires them to apply their learnings to create strong and memorable passwords that satisfy all the heuristics.

3.1.2. Game Mechanics

The game storyline is based in a fictional world. Learning experiences have found to be enhanced using story-based agents [41]. The gameplay of a level starts with a pre-test, then the game, followed by a post-test. The game consists of two levels, Level 1 and Level 2, and each level consists of different sequential stages for pre-test, gameplay, password creation, distraction task, password recall, and post-test (cf. Figure 1). Based on our design choices, we framed our game story and various gameplay elements, as follows:

Jungle environment: The protagonist, Soma, is an archaeologist who is in search of two ancient artifacts that were lost years ago in a land called "Passworld". Soma has to travel two days and two nights through the jungles of Passworld to find them (cf. Figure 2). The two days are represented by Level 1 and Level 2 gameplays.

Securely storing the artifact: Since the ancient artifacts are precious, Soma has to store them after collecting, to protect them from being stolen. This is done by creating strong gates (analogous to secure passwords) around Soma's camp. This happens in the two password creation stages (cf. Figure 4), represented by two nights.

Learning the password heuristics: In the two main levels, the players can interact with oncoming animals during gameplay to learn about various password heuristics (cf. Figure 3). Every oncoming animal will raise the curiosity of the player by showing basic heuristic details as a riddle (E.g. Fox in. Figure 2). The player can choose to "*know more*" about a particular heuristic by clicking the animal's heuristic text. This will pop up a detailed description of the heuristic and certain statistics associated with it (cf. Figure 3). If these heuristics are not satisfied during each password creation stage, the corresponding animal will attack the password gate (cf. Figure 5), and enter the camp. This also signifies how a password meter checks for various heuristics [48]. While in [48], the users are not required to satisfy all the password heuristics in an entered password, our game has this requirement as we wanted to teach all the available heuristics to the users and tell them that every single one of them is important.

Resource gathering: The resources for creating these gates are obtained throughout the journey, in the form of tablets with character classes mentioned as L, U, D, and S (Lowercase, Uppercase characters, Digits, and Symbols) (cf. Figure 7). In real life, these character classes are required for creating a password.

Creation of password gates: Once the player collects the artifacts, stores them using secure password gates (cf. Figure 4), they complete one full day in the game. We introduced two activities post each password creation stage that act as distraction tasks. Distraction tasks [31] distract the players for a brief period after password entry, to encourage them to create memorable passwords. Our tasks are two mini activities that ask the players to arrange certain items correctly (using drag and drop) to a) Ignite a campfire b) Cook food (cf. Figure 8). This step is added to promote awareness about the importance of creating memorable passwords. To continue the journey further on the next day, the player has to unlock the gate using the same password (cf. Figure 4). This password recall stage is where password memorability is tested.

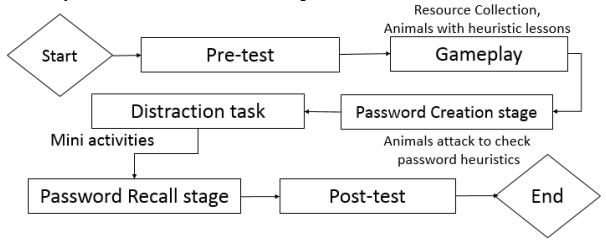

Figure 1: Passworld Game Level Flow

If players fail a level, it can be replayed again. The game did not have timers as these might have created unwanted sense of urgency that could have limited gameplay experience [43]. Passworld used simple controls with arrow keys for navigation, jumping, and mouse clicks for selections.

3.2. Implementation of Game-based Training

The game implements the use of feedback and instructions to promote learning throughout.

Figure 2: Passworld Level 1 Gameplay

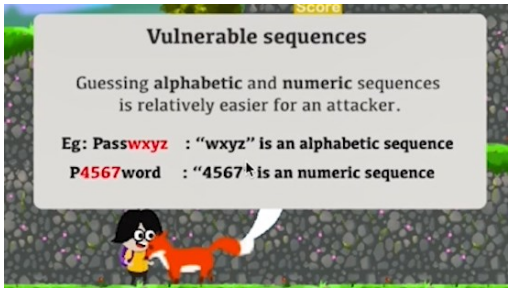

Figure 3: Heuristic description received from animal

Figure 4: Password creation (Gate Creation) stage. The password recall stage is similar in appearance, but without the options to view password heuristics (yellow animal icons).

Figure 5: Animal (fox) approaching to attack the gate.

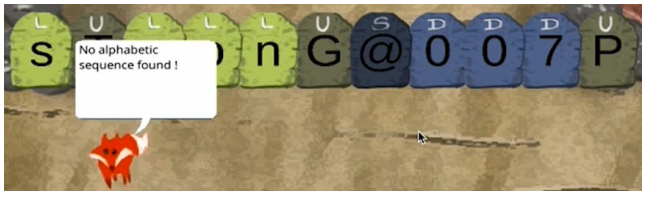

Figure 6: Animal returning as the corresponding heuristic (no alphabetic sequences found) was satisfied.

Figure 7: Game Resources and artifacts- L, U, D, S; the numbers on the tablets correspond to the count of each collected resource. The artifacts (one for each Level) depict important information that the player should store securely.

Figure 8: Distraction Task, Level 2

The feedback methodology was added to promote self-learning and reflection [**15**]. The game implements the following methods to promote learning.

Instructions: The oncoming animals provide instructions on various heuristics to the players. We added different animals to provide a visual identity to each heuristic, to make it more memorable. The same heuristic information is also available during password creation stages to help users learn the password creation strategies. Therefore, players who feel reading the information during gameplay is disruptive can read it at a further time, while creating the password. The heuristics are taught one at a time, but at the end of the level, the created password should incorporate all these heuristics. Even though a strong password does not need all these heuristics to be satisfied [48], we did this in order to teach and make the users understand that every heuristic is important.

Feedback: As soon as the player enters a password during the password creation stage, immediate feedback is received, indicating to them the potential vulnerabilities within the password entered. The study on adaptive password blacklisting policies [38] introduces an interface to provide the users with suggestions on modifying passwords to conform to the policies. We used our study to make this process voluntary. We did not provide suggestions to the users, but only feedback on if they satisfied certain password heuristics or not.

3.3. Game Data Recorded

Data captured and stored in the form of game data included the demographic information, pre-test, post-test, and feedback survey responses, various time stamps, gameplay

Password Heuristics	Heuristic Description (and Corresponding Animal in Game)	Levels taught	Identi-fier	Clus-ter
Length	Password length must be more than 8 **(Default policy)**	1,2	H1	
Lowercase present	At least 1 lowercase character must be present **(Default policy)**	1,2	H2	
Uppercase present	At least 1 uppercase character must be present **(Default policy)**	1,2	H3	C1
Digit present	At least 1 digit must be present **(Default policy)**	1,2	H4	
Symbol present	At least 1 symbol must be present **(Default policy)**	1,2	H5	
Repeated characters	Not more than 3 consecutive repeating characters (e.g. "honeeeey" has 4 repeating 'e's,). **(Animal: Raccoon)**	1,2	H6	
Duplicated characters	Number of duplicated characters should not be more than 50% of total character count in password **(Animal: Porcupine)**	1,2	H7	
Repeated sections	3 or more repeating set of characters (e.g. honeyhoney honey, honeyyenohhoney) should not be present **(Animal: Fox)**	1,2	H8	C2
Alphabetic sequences	Not more than 3 consecutive Alphabetic or numerical sequences should be present(e.g. 12345, ghiJkLm etc.) **(Animal: Fox)**	1,2	H9	
Predictable positions of:	Checking the predictable positions of symbols, digits and uppercase characters **(Animal: Monkey)**	2	H11	
Symbols	Symbols should not be present just at the end (e.g. passwor@) or password should not be having a common "letters-symbols-digits" format (e.g. pasS@921)			
Digits	Digits present in the beginning or at the end or password having all digits is a common pattern, should be avoided	2	H12	
Uppercase Characters	UPPERCASE order present in the beginning or all characters in password being uppercase, is a common pattern, should be avoided	2	H13	C3
Predictable structure	Password should not fall under a set of common password structures (e.g. LSLLLDLD) **(Animal: Leopard)**	2	H14	
Keyboard patterns	4+ equally spaced keys in password (e.g. QWERTY) **(Animal: Hyena)** should not be present (US-English language keyboards).	2	H10	
Date formats	DDMMYY – with and without delimiters and years (1900-2049), in all formats including month names should not be present. **(Animal: Bear)**	2	H15	C4
Blacklists	Password should not contain common organization related words anywhere in it. **(Animal: Snake)**	1,2	H16	C5

Table 1: Password Heuristics Checked Within the Game

data, heuristics viewed by the player, password structures entered, level attempt counts, heuristics (failed and successful) and password creation, recall attempts. The game converts passwords entered by users into their respective character structures and stores in the database. Password structure [38] is an ordered sequence that captures the password's composition using four character classes. These classes are L, U, D, and S, for lowercase and uppercase characters, digits, and symbols respectively. For example, a password like "P@ssw0rd" will only be stored as "USLLLDLL" instead of its plain text for analysis.

3.4. Password Heuristics

The game trains in a set of 16 password heuristics, with each heuristic being tagged to a particular animal (as shown in Table 1). A previous study by Ur, et al. [48] found them to be effective in increasing password strength. These password heuristics, categorized as two sets based on increasing complexity, were added to the two game levels. The first level has basic password requirements like length (H1), presence of character classes (H2-H5), alphabetic sequences (H9) etc. of which, H1-H5 were part of our organization's default password policies. The second level focuses on the heuristics from the first level along with new heuristics that check formatting, repeated sections in passwords, date formats etc. A set of common words related to the organization (classified as "blacklisted" passwords) were added as a check as well. We also compared user created password structures with over 2,124 structures obtained from the previous study [48]. These

heuristics were also clustered based on the common characteristics they possess, as C1 having basic password heuristics, C2 with character sequences, C3 with predictable positions, C4 having certain patterns, and C5 with the blacklists. We have taught these heuristics through the game's main levels, and let the players incorporate these heuristics while creating the passwords in password creation stages, thus letting the players demonstrate what they have learnt.

3.4.1. How User Created Passwords are Checked

The game checks users' passwords through the following steps:

a. As soon as a player uses the resources (L, U, D, and S) to form a password gate, default checks for length and presence of all character classes are done. If any of them is not satisfied, the game shows appropriate error messages to the player instantly.

b. Once the password satisfies the basic criteria, the password heuristics evaluation begins.

c. For each heuristic, an animal approaches the gate (cf. Figure 5). If the corresponding heuristic is satisfied, the animal leaves (cf. Figure 6); else, it attacks the gate and enters the camp resulting in a penalty as loss of life. This process repeats until all level heuristics are satisfied (level cleared) or when all life is lost (level failed). After this, the player continues to the next level or goes back to the start of the level respectively.

4. Study Design

The goals of our study were to find the effectiveness of a game-based enterprise password awareness training on various password heuristics, and to identify if such a training could be beneficial to enterprise password diversity. Previous studies [8, 41] show that game-based methods have shown better results than text-based means, when it comes to cybersecurity training. We utilized this result to test if games could help in password awareness training, and we measured this using pre and post-tests along with the game. The following sections show our study procedure and evaluation results.

4.1. Participant Demographics

Game participants were employees of our organization. They were recruited for the study using mailers about the game. Interested participants clicked on the game URL within the mail to access the game. Though equipped with computer knowledge, the participants had varied understanding of gaming and password awareness. Passworld was online for one month and was played by 4,906 participants from around the globe. We selected a set of lucky winners from the participants who completed the game (20 people per day), and rewarded each of them using our organization's equivalent of virtual currency (with a monetary value of approximately $4). Table 2 shows the demographics data collected from the participants.

Criteria	Percentage (%)
Gender	
Female	43.15
Male	53.75
Others	0.12
No answer	2.98
CS/IT Education	
No	36.44
Yes	60.01
No answer	3.55
Age Group	
21-30	73.44
31-40	20.08
41-50	2.83
Above 50	0.45
No answer	3.20
Educational Degree	
Undergraduate Degree	75.30
Master's Degree	20.12
Doctorate	0.22
Others	1.22
No answer	3.14

Table 2: Participant Demographics

4.2. Procedure

We organized the study as a three-step methodology. Initially the participants had to answer a pre-test (step 1). This was followed by the actual gameplay (step 2) and then the post-test (step 3). The participants accessed the game using their respective devices and those participants who completed all the game levels from beginning to end were included in our evaluation. Only the first successful attempt of completion was used in our data analysis, even though many participants returned to play the game more than once. We measured the attempt count by tracking the "participant id" of participants, which was assigned based on their hashed email addresses. Evaluation on users' password knowledge improvement was done by analyzing their responses to the pre- and post-tests, and the password

structures entered by them during the password creation stage.

Each test question covered a password heuristic. The pre- and post-test questions followed a similar format asking players to select the relatively weaker password between two given choices. We created the password choices by picking suitable passwords from leaked databases [1, 58] and minimally modifying them to be able to test a particular heuristic, similar to the method followed in [50]. For example, to test H13 (predictable position of uppercase character) we chose the password "*brooklyn*" from the leaked database [58], and created the password pair comprising of "*Brooklyn*" and "*brooklYn*", of which the former is weaker as the uppercase character is at a very predictable position. This method was extended to the password pairs of other questions as well. Participants were also asked to provide their confidence ratings for every answer. The survey questions are provided as appendix.

5. Performance Evaluation

We evaluated the impact of our game-based training on users' password creation strategies. We were also interested in the changes in users' knowledge levels when it came to password practices by measuring correct answers given for pre and post-tests. We tried to answer our initial research questions through the study.

5.1. Can a Game-based Training Teach Password Heuristics to Participants?

We measured the effects of the game by analyzing the improvement in correct answers and confidence levels of the participants. The results are detailed in the following sections.

5.1.1. Participants' Correct Answers

We analyzed the participants' pre and post-level test responses. The test questions prompted users to choose the weaker between two given passwords. Passwords were shown such that the pairs focused on one heuristic each, and the password with the absence of that particular heuristic was considered weaker in that pair.

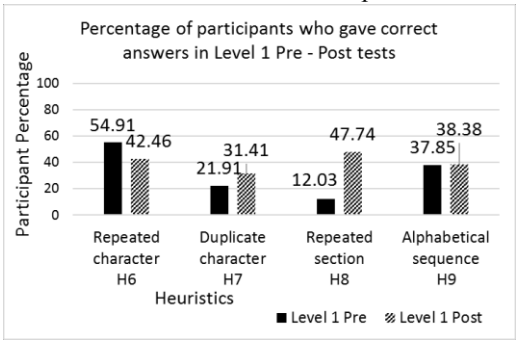

Figure 9: Level 1 - Participants' Correct Answers

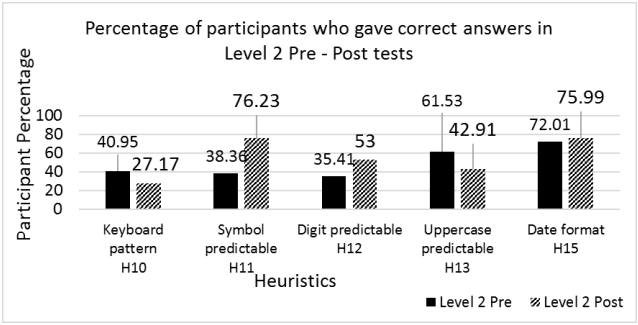

Figure 10: Level 2 - Participants' Correct Answers

The results show that more participants were able to provide correct answers to test questions after playing the game compared to before. There are however, exceptions for heuristics H6, H10, H13. The plots (cf. Figures 9, 10) show the participants' performance based on password heuristics in both the levels. H6 had a reduction of 12.45%, H10 had 13.78%, and H13 had 18.62% reduction in correct answers in post-level tests compared to pre-level tests. While we are unable to determine the cause for reduction in H6 now, the reduction in H10 could be attributed to the large possible combination of patterns possible.

The variety of keyboard patterns might turn out to be difficult for users to remember, which could mean that active feedback during password entry might be better for such heuristics. The predictable positions of uppercase characters (H13) had a decrease probably because the game time was not sufficient in unlearning the common practices of adding uppercase letters at the beginning or end. This may require multiple training sessions to unlearn. The users did show improvements in Level 2 while entering passwords by avoiding uppercase characters at the beginning. Questions on H1-H5 were not asked during the tests as these belonged to basic password requirements of our organization that the employees were aware of. Combining both levels, we asked the players 24 questions, 12 each in pre and post-level tests. The average number of correct answers increased from 5.96 (pre-test, *SD=2.3*) to 6.57 (post-test, *SD=2.69*). A statistically significant difference was observed with respect to the correct answers given by participants in the tests before and after the game (two-tailed paired t-test, $t(4905) = -19.35$, $p < .001$). Questions on H14 (predictable structures) were not asked separately in pre and post-tests as these would also fall under other heuristics. H14 was tested during password creation level; however, the passwords that satisfied the heuristics H1-H13 also satisfied H14. Averaging correct answer percentages of certain clusters (for both tests), we found an 8.32% increase in C2, 12.28% increase in C3, and 4.9% decrease in C4 (cf. Figure 11), showing that patterns

(especially keyboard patterns) is where participants need to gain more knowledge on.

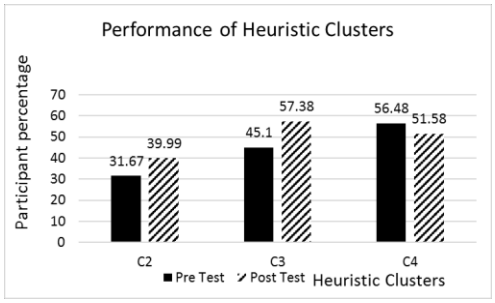

Figure 11: Performance of Heuristic Clusters

The previous study on GAP game [47] that tested password practices, similar to heuristics H11, H12, and H13, had shown better results with an average of 80.17%, 78.45%, and 87.93% participants correctly identifying password practices, H11, H12, and H13 respectively. Our game had a larger set of heuristics, with Level 1 requiring users to check ten heuristics and Level 2 having an additional six heuristics apart from Level 1 heuristics. This might have caused an overload of learning content in one go, as participants have shown better results for H13 in GAP [47] compared to Passworld. This could also show that while learning these heuristics is important, the manner to train users in them could be gradual. The users could be trained in an initial set of heuristics at first, and afterwards, the next set of training could be undergone.

5.1.2. Participants' Confidence Results

Participants were asked to rate their confidence level for their answers in both the tests. Each question had five levels of confidence (1: Not confident at all (Least Confident), 2: Not confident, 3: Neutral, 4: Confident, 5: Very confident). From Figures 12 and 13, we can see that the confidence level of players had a consistent increase. The average confidence rating increased from 4.39 (pre-test, variance =0.45) to 4.47 (post-test, variance =0.44). The analysis of the confidence levels shows the results to be statistically significant (Wilcoxon signed-rank test, $z = -18.87$, $p < .0001$).

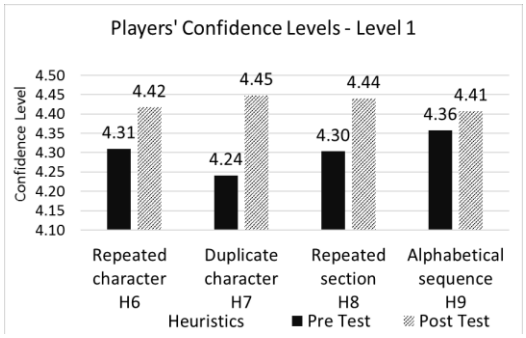

Figure 12: Level 1 - Participants' Confidence Levels

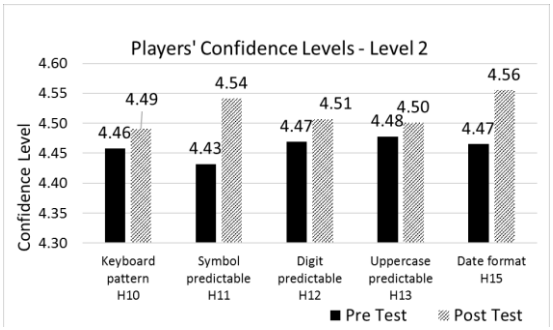

Figure 13: Level 2 - Participants' Confidence Levels

Considering confidence levels of clusters, average confidence level of C2 increased from 4.30 to 4.42, C3 increased from 4.45 to 4.51, and C3 increased from 4.46 to 4.52 (cf. Figure 14).

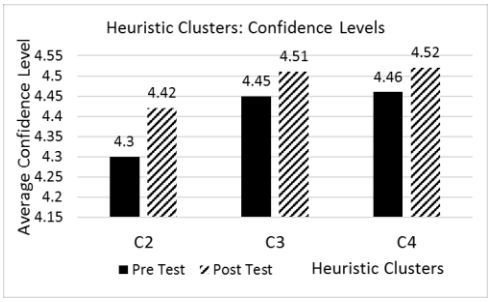

Figure 14: Confidence Levels of Heuristic Clusters

5.1.3 Demographic Analysis

Among the participants, the age group 21-30, and Bachelor's Degree holders showed a consistent increase in almost all heuristics. People with non-IT background (H9: 20% increase, H11: 104%) showed more improvements than the participants with IT background (H9: 5% increase, H11: 96%). H8 has better improvements in all demographics (with highest for Master's degree holders, 373% from pre to post). We also conducted a comparative study on gender differences vs. password practices. Petrie, et Al [36], also suggested a similar study. From our analysis, it was found that women did a better job of recalling created passwords, with 78.05% (Level 1) and 86.06% (Level 2) recalls matching their created passwords compared to men with 72.65% (Level 1) and 81.98% (Level 2). This supports the behavior reported in the study [36] where men expressed greater difficulty in remembering passwords.

5.1.4. Failure Count of Heuristics

Often, participants took several attempts before creating a password that satisfied each Level's heuristics. Counting all such attempts, for H1-H9 and H16, we calculated the overall number of times the participants failed for each heuristic before satisfying them. These heuristics appear in both levels 1 and 2. From Figure 15, we can see that the total number of failed attempts for each heuristic has come

down drastically from Level 1(*M= 283*) to Level 2 (*M= 106*). We can infer that once participants gained sufficient knowledge through game-play on various common password practices, they incorporated these learnings during password creation, thereby satisfying the heuristic checks. The number of failed attempts for blacklisted passwords decreased from 845 in Level 1 to 195 in Level 2 (~77% reduction), showing improvement. Alphabetical sequences (H9) however, showed a slight increase in failure rates. This could be attributed to the fact that there are a large number of possible sequences to avoid, and the participants were having difficulty in identifying them.

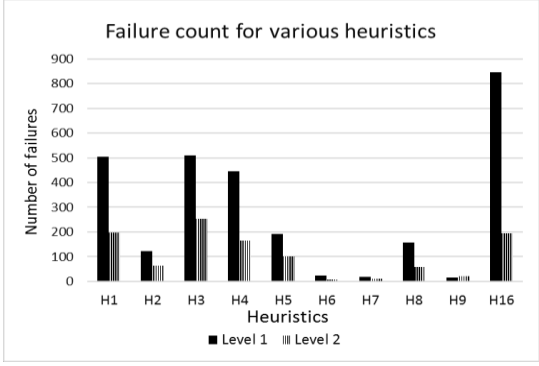

Figure 15: Heuristics vs Total Failures by Participants

5.2. Can Such a Training on Heuristics Improve the Participants' Password Diversity?

From the user responses, we found changes in password distribution when more heuristics were satisfied. We analyzed this using password structures. An interesting observation was about the use of common password structures. In the password creation stages, several participants followed similar structures while creating passwords (Table 3, Column 1). This shows that most of the players' ideas on secure passwords might end up being wrong, as shown in a previous study [49].

5.2.1. Level-wise Differences in Password Structures

Comparing the top 10 most popular password structures from Level 1 and Level 2, we were able to identify a major change in the number of occurrences of similar structures. Table 3 shows the comparison, where structures with common patterns (ULLLLSDDDD, for instance could be Abcde@1234) have shown drastic decrease in the number of occurrences. Level 2 structures are more spread out and with different password patterns, which suggest that participants created diverse passwords while satisfying more heuristics. It also brought down many common password structures, like those with common numerical sequences DDDD (e.g. 1234), uppercase characters (U) only in the beginning, etc.

Popular Password Structures with their Occurrences

Level 1	Level 2
ULLLLSDDDD(84)	LSLLLDLU(52)
ULLLSDDDD(64)	LSLLLDLLU(20)
ULLSDDDD(56)	LLLUSDLLL(10)
ULLLLLSDDDD(52)	ULLSULDDLL(8)
ULLLLLSDD(46)	LLLUSDDDUU(8)
LSLLLDLU(45)	UUUUSDDLL(6)
ULLLSDDD(44)	USLSUDLU(6)
ULLLLLLSDD(39)	SLLDLLLU(6)
ULLLLSDD(37)	LULLLSULLDDDL(5)
ULLLLSDDD(35)	ULSULDDDDL(5)

Table 3: Differences in occurrences of popular password structures (and their occurrences) from Level 1 compared to Level 2, after the initial analysis.

In Level 1, majority of the passwords start with an uppercase character, and the character does not appear anywhere else in the password. The game teaches this practice to be less secure, as one of the heuristics (H13) pointed out that use of uppercase in the beginning is a common practice. In the process of trying to satisfy more password heuristics into their structures, this trend was reduced. From the participant data, we found that the participants created 17,319 passwords that fell in 11,286 different structures. Considering only the password structures that satisfied the heuristics (4,906 in each level), there were 3,595 and 4,451 different structures respectively in Level 1 and Level 2.

Furthermore, 3,246 (66.16%) and 4,182 (85.24%) password structures were unique in Level 1 and Level 2 respectively, having only one occurrence. This trend showed that when more heuristics were to be satisfied in passwords, enterprise users could create passwords that were more diverse.

5.3. Other Game Data

For the password recall stage, data shows that 55.54% players matched the passwords in their first attempt for Level 1 and it increased to 71.70% for Level 2. Considering the players who recalled their passwords, in multiple attempts, the numbers come to 3,700 (75.41%) for Level 1 and 4,122 (84.01%) for Level 2. The participants who successfully recalled passwords for both levels comes to 3247(66.18%).

5.3.1. Player Involvement

Overall, 6,814 participants showed interest in playing the game, of which, 4,906 completed it. A decrease of 28%

could be because of network and proxy issues in certain locations, as noted from the participant feedback comments. Among the 4,906 participants, we calculated the player involvement by measuring number of resources and artifacts collected per level by the players (Table 4). The data shows that majority of players have had an immersive gameplay, with increasing number of collected artifacts and resources nearing the total number available. Here L, U, D, S correspond to character classes and A denotes the artifact collected.

		L	U	D	S	A
Level 1	Average	5.37	7.05	6.7	3.64	0.72
	Total	6	9	9	4	1
Level 2	Average	11.4	7.48	6.54	11.05	0.83
	Total	12	8	7	12	1

Table 4: Gameplay Resources collection data

5.4. Player Feedback

The general feedback by participants included comments like "informative and interesting", "Excellent game to deliver the message on the usage of strong passwords", "Fun and creative game. Nice idea." etc. While some participants requested for the game to be available as a permanent training method, some participants found the game time of over 15 minutes to be a bit longer. At the end of the post-test, we asked the participants to rate Passworld on a 5-point Likert scale [22] with respect to fun, education, and learning. As per the feedback data (cf. Figure 16), 93.50% participants agreed the game to be fun (*M = 4.42*, *SD = 0.69*), 93.85% (*M= 4.42, SD= 0.68*) found the game to be educational and 94.24% (*M= 4.48, SD = 0.68*) considered they have learned about secure password practices.

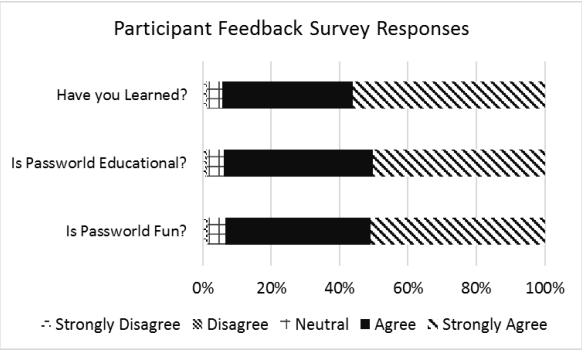

Figure 16: Participants' Feedback Data

6. Discussion

The Passworld game was successfully completed by 4,906 employees in our organization. The analysis of their gameplay data found the game to be effective and engaging. It also helped us find various insights on enterprise users' password practices. The insights are as follows.

a) Participants showed improvements in creating passwords that satisfied various heuristics like avoiding date formats, predictable positions of symbols and digits, character duplicates to name a few. Their new passwords satisfied more number of heuristics after users played the game levels, compared to before.

b) Prior to satisfying all game heuristics, participants created similar password structures, as seen from Column 1 of Table 3. After satisfying more heuristics, the participants were able to create more diverse passwords. Over 90% of passwords created after satisfying all the game heuristics resulted in unique password structures. This trend, when followed outside the game, would result in password diversity across organization.

c) It is to be noted that the diversity in user-created password structures post gameplay is not brought about by providing suggestions on modifications to the entered password, unlike in the previous study [38]. More likely, the users applied their learnings obtained from the game to satisfy more heuristics in their passwords. This resulted in more password diversity.

d) Comparing with the previous game GAP [47] which showed good results with a small set of heuristics, we found that certain heuristics like repeated characters (H6), keyboard patterns (H10), and predictable uppercase characters (H13) showed decrease in performance after game-play. H10 can be considered as a complex heuristic with several possible patterns that the user has to know of. H13 is such that addition of uppercase characters at the beginning of a password is a very common practice that might need multiple trainings to unlearn. This could suggest that these heuristics might either need a more in-depth training or an active feedback while passwords are being entered in real time. The results could also mean that training on a very large set of heuristics might have caused a learning overload that might have resulted in decrease in results in Level 2. This could be avoided by providing a gradual training on heuristics.

e) Considering the heuristic clusters, C4 had a decrease of around 5% showing that "identifying patterns" is where the participants struggled the most. Real-time suggestion to avoid keyboard patterns during password entry could be a method to reduce the users from including patterns into their passwords.

f) We received certain insights on changes to gameplay that include reducing the overall game length, reducing the time needed for animals to attack the gate, and distribution of heuristics into various levels.

6.1 Limitations

Our study has a certain number of limitations that have to be considered while analyzing the results. The participant demographics selected for the game included employees of our organization, who may know basics of cybersecurity or password practices. The game data of 4,906 participants shows positive responses in understanding password heuristics and creation strategies. However, to test the effectiveness of the game in a real-world scenario, we will have to monitor individual password practices throughout the organization to start with. This process could violate the privacy of individuals, and therefore has not been attempted by us. The results of a previous research [28] shows that the positive results of a smaller lab experiment were carried forward to a larger audience. We hope that our experimental results show a similar trend. Similarly, the passwords entered by the users within the game were most likely not their real passwords used in any accounts. While a research study [16] suggests that the passwords implemented by users in a study could resemble real life passwords, the extent to which our participants would utilize similar passwords in both game and real life needs to be studied. An alternate form of experiment, e.g. a text-based condition or another control condition, has not been tested. While the study makes no claims of being better than a controlled experiment, our training shows that a fun oriented gameplay can help teach password heuristics to users. More studies need to be conducted in order to obtain conclusive evidence on the effectiveness of a game as a medium to train enterprise users on password awareness.

The game makes no claims about making all the user passwords memorable. It attempts to improve the understanding of people in creating stronger passwords than what they used to create prior to the game. While we observed reasonable password recall rates in our study, we do not have conclusive evidence on password memorability over long periods.

Passworld tries to offer a starting point in the area of game-based password security education. A previous research study [39] show that user behavior related to password usage can be influenced with positive reinforcement. The results of our study suggest that game-based training could also influence users' behaviors related to passwords. A solid understanding on password practices coupled with the use of a password meter could provide better security in terms of password strength.

7. Conclusions

The Passworld game was designed to provide awareness on various password heuristics to enterprise users. The main objectives of our study were to find 1) if a game-based training could teach users on password heuristics 2) if such a training on heuristics could improve organizational password diversity. We used the password heuristics from a previous study [48] for teaching, and we checked if the users satisfied every one of these heuristics during their password creation. Our intention was different from the previous study [48], where satisfying all the heuristics was not mandatory. We intended to teach the users about the importance of each heuristic, and wanted to see how many users successfully implemented what they learnt. We presented the results from our enterprise study with 4,906 participants. Even though our study was a standalone study, without a control condition, we found that after playing the game, the correctness and confidence levels of the participants have increased. The password structures created by the participants have shown more diversity post gameplay. This, along with the positive feedback, shows that the gameplay has helped the participants learn the concepts to implement diverse and memorable passwords. We believe that this trend, when followed in real life, would result in organizational password diversity.

We recommend the launch of such training methods in an organizational environment to ensure that users learn about password heuristics and incorporate them while creating passwords to promote diversity in password structures. This could be a deciding factor when it comes to organizational password security.

Recommendations for further study follow. First, the study could be carried out on a set of wider demographics, with different levels of understanding of security concepts and learning backgrounds. To reduce the information overload, we propose a gradual learning with one set of heuristics, followed by another set. Training on patterns (like alphabetic sequences, keyboard patterns etc.) could be done separately to lay emphasis on it. A methodology to evaluate users' password memorability over long periods could also be beneficial in proposing further learning goals. We aim to explore further areas of password and cybersecurity education through interactive gameplay experiences.

References

[1] A. Vance. 2010. "If your password is 123456, just make it hackme," New York Times, Retrieved February 19, 2020 from https://www.nytimes.com/2010/01/21/technology/21password.html.

[2] Adams, A., and Sasse, M.A. User are not the enemy. Commun. ACM 42, 12 (Dec. 1999), 40–46

[3] Aloul, Fadi, Syed Zahidi,l Wassim El-Hajj. "Two factor authentication using mobile phones." 2009 IEEE/ACS International Conference on Computer Systems and Applications. IEEE, 2009.

[4] Annetta, Leonard A. "The "I's" have it: A framework for serious educational game design." Review of General Psychology 14.2 (2010): 105.

[5] Bellotti, F., Ott, M., Arnab, S., Berta, R., de Freitas, S., Kiili, K. and De Gloria, A., 2011, October. Designing serious games for education: from pedagogical principles to game mechanisms. In Proceedings of the 5th European Conference on Games Based Learning. University of Athens, Greece (pp. 26-34).

[6] Buttyán, Levente. "Introduction to IT security." 2016.

[7] Cheung, Wing-Yee, et al. "Back to the future: Nostalgia increases optimism." Personality and Social Psychology Bulletin 39.11 (2013): 1484-1496.

[8] CJ, G., Pandit, S., Vaddepalli, S., Tupsamudre, H., Banahatti, V., & Lodha, S. (2018, October). Phishy-a serious game to train enterprise users on phishing awareness. In Proceedings of the 2018 Annual Symposium on Computer-Human Interaction in Play Companion Extended Abstracts (pp. 169-181).

[9] Compton, Kate, and Michael Mateas. "Procedural Level Design for Platform Games." AIIDE. 2006.

[10] Cone, Benjamin D., Cynthia E. Irvine, Michael F. Thompson, and Thuy D. Nguyen. "A video game for cyber security training and awareness." computers & security 26, no. 1 (2007): 63-72.

[11] Csikszentmihalyi, Mihaly, Sami Abuhamdeh, and Jeanne Nakamura. "Flow." Flow and the foundations of positive psychology. Springer, Dordrecht, 2014. 227-238.

[12] Daitch, Heidi. 2017. 2017 Data Breaches- The Worst So Far. Retrieved February 19, 2020 from https://www.identityforce.com/blog/2017-data-breaches

[13] De Carnavalet, Xavier De Carné, and Mohammad Mannan. "From Very Weak to Very Strong: Analyzing Password-Strength Meters." NDSS.Vol. 14. 2014.

[14] Denning, Tamara, Adam Lerner, Adam Shostack, and Tadayoshi Kohno. "Control-Alt-Hack: the design and evaluation of a card game for computer security awareness and education." In Proceedings of the 2013 ACM SIGSAC conference on Computer & communications security, pp. 915-928. ACM, 2013.

[15] Donovan, M. Suzanne, John D. Bransford, and James W. Pellegrino. How people learn: Bridging research and practice. National Academy Press, 2101 Constitution Avenue NW, Lockbox 285, Washington, DC 20055, 1999.

[16] Fahl, S., Harbach, M., Acar, Y., and Smith, M. On the ecological validity of a password study. In Proc. SOUPS (2013)

[17] GameDev.Net.2017. Intrinsic and Extrinsic Motivation. Retrieved May 29, 2020 from https://www.gamedev.net/tutorials/game-design/game-design-and-theory/intrinsic-and-extrinsic-motivation-r4713/

[18] Giannakas, Filippos, Georgios Kambourakis, and Stefanos Gritzalis. "Cyberaware: A mobile game-based app for cybersecurity education and awareness." Interactive Mobile Communication Technologies and Learning (IMCL), 2015 International Conference on. IEEE. 2015.

[19] Gordon, William J., Adam Fairhall, and Adam Landman. "Threats to Information Security—Public Health Implications." New England Journal of Medicine 377.8 (2017): 707-709.

[20] Habib, Hana, Pardis Emami Naeini, Summer Devlin, Maggie Oates, Chelse Swoopes, Lujo Bauer, Nicolas Christin, and Lorrie Faith Cranor. "User behaviors and attitudes under password expiration policies." In Fourteenth Symposium on Usable Privacy and Security ({SOUPS} 2018), pp. 13-30. 2018.

[21] Hudson Soft.1986.Adventure Island. Game [MSX, NES] (September 12, 1986) Hudson Soft., Tokyo, Japan.

[22] I. Elaine Allen and Christopher A. Seaman. 2007. Likert Scales and Data Analyses. July 2007. Retrieved February 19, 2020 from http://rube.asq.org/quality-progress/2007/07/statistics/likert-scales-and-data-analyses.html

[23] Ives, Blake, Kenneth R. Walsh, and Helmut Schneider. "The domino effect of password reuse." Communications of the ACM 47.4 (2004): 75-78.

[24] Jay Jay. 2018. Despite increased cyber-risk awareness, poor password hygiene still rules. Retrieved February 19, 2020 from https://www.scmagazineuk.com/despite-increased-cyber-risk-awareness-poor-password-hygiene-rules/article/1472772

[25] Johnson, B. R., and Koedinger, K. R. 2002. Comparing instructional strategies for integrating conceptual and procedural knowledge. In Proceedings of the Annual Meeting [of the] North American Chapter of the International Group for the Psychology of Mathematics Education, vol. 1–4, pp. 969–978.

[26] Kiili, Kristian. "Digital game-based learning: Towards an experiential gaming model." Internet and Higher Education 8 (2005): 13-24.

[27] Krathwohl, David R. "A revision of Bloom's taxonomy: An overview." Theory into practice 41.4 (2002): 212-218.

[28] Kumaraguru, Ponnurangam, et al. "Teaching Johnny not to fall for phish." ACM Transactions on Internet Technology (TOIT) 10.2 (2010): 1-31.

[29] Kumaraguru, Ponnurangam, Yong Rhee, Alessandro Acquisti, Lorrie Faith Cranor, Jason Hong, and Elizabeth Nunge. "Protecting people from phishing: the design and evaluation of an embedded training email system." In Proceedings of the SIGCHI conference on Human factors in computing systems, pp. 905-914. ACM, 2007.

[30] Lee Mathews. 2017. File With 1.4 Billion Hacked And Leaked Passwords Found On The Dark Web. Retrieved February 19, 2020 from https://www.forbes.com/sites/leemathews/2017/12/11/billion-hacked-passwords-dark-web/#516d814221f2

[31] Liu, Ye, and Xiaolan Fu. "How does distraction task influence the interaction of working memory and long-term memory?." International Conference on Engineering Psychology and Cognitive Ergonomics. Springer, Berlin, Heidelberg, 2007.

[32] Moreno R, Mayer RE, Spires HA, Lester JC. The case for social agency in computer-based teaching: Do students learn more deeply when they interact with animated pedagogical agents? Cognition and instruction. 2001 Jun 1;19(2):177-213

[33] New York Film Academy.2015. Learning From The Best: Action-Adventure Games. Retrieved February 19,2020 from https://www.nyfa.edu/student-resources/learning-from-the-best-action-adventure-games/

[34] Nintendo.1985. Super Mario Bros. Game [NES] (September 13, 1985) Nintendo, Kyoto, Japan. Level/area: World 8-4

[35] Ofcom. Adults media use and attitudes. Technical report, Ofcom: UK Communications Regulator, 2016. https://www.ofcom.org.uk/__data/assets/pdf_file/0026/80828/2016-adults-media-use-and-attitudes.pdf (accessed January 2020).

[36] Petrie, Helen, and Burak Merdenyan. "Cultural and Gender Differences in Password Behaviors: Evidence from China, Turkey and the UK." Proceedings of the 9th Nordic Conference on Human-Computer Interaction. ACM, 2016.

[37] Rass, Stefan, and Sandra König. "Password Security as a Game of Entropies." Entropy 20.5 (2018): 312.

[38] Segreti, Sean M., William Melicher, Saranga Komanduri, Darya Melicher, Richard Shay, Blase Ur, Lujo Bauer, Nicolas Christin, Lorrie Faith Cranor, and Michelle L. Mazurek. "Diversify to survive: making passwords stronger with adaptive policies." In Thirteenth Symposium on Usable Privacy and Security ({SOUPS} 2017), pp. 1-12. 2017.

[39] Seitz, Tobias, and Heinrich Hussmann. "PASDJO: quantifying password strength perceptions with an online game." Proceedings of the 29th Australian Conference on Computer-Human Interaction. ACM, 2017.

[40] Selvarajan, Balamurugan. "Simple two-factor authentication." U.S. Patent Application No. 11/267,148.

[41] Sheng, Steve, Bryant Magnien, Ponnurangam Kumaraguru, Alessandro Acquisti, Lorrie Faith Cranor, Jason Hong, and Elizabeth Nunge. "Anti-phishing phil: the design and evaluation of a game that teaches people not to fall for phish." In Proceedings of the 3rd symposium on Usable privacy and security, pp. 88-99. ACM, 2007.

[42] Sherry, John, and Angela Pacheco. "Matching computer game genres to educational outcomes." Electronic Journal of Communication/La Revue Electronique de Communication16.1& (2006): 2.

[43] Sweetser, P., & Wyeth,P.(2005). GameFlow: a model for evaluating player enjoyment in games. Computers in Entertainment (CIE), 3(3), 3-3.

[44] Thompson, Michael, and Cynthia Irvine. "Active learning with the CyberCIEGE video game." (2011)

[45] Bradley, Tony. 2018. Security Experts Weigh In On Massive Data Breach Of 150 Million MyFitnessPal.Accounts Retrieved February 19, 2020 from https://www.forbes.com/sites/tonybradley/2018/03/30/security-experts-weigh-in-on-massive-data-breach-of-150-million-myfitnesspal-accounts/#209ddebd3bba

[46] Hunt, Troy. 2019. The 773 Million Record "Collection #1" Data Breach. Retrieved February 19, 2020 from https://www.troyhunt.com/the-773-million-record-collection-1-data-reach/

[47] Tupsamudre, Harshal, Rahul Wasnik, Shubhankar Biswas, Sankalp Pandit, Sukanya Vaddepalli, Aishwarya Shinde, C. J. Gokul, Vijayanand Banahatti, and Sachin Lodha. "GAP: A Game for Improving Awareness About Passwords." In Joint International Conference on Serious Games, pp. 66-78. Springer, Cham, 2018.

[48] Ur, Blase, Felicia Alfieri, Maung Aung, Lujo Bauer, Nicolas Christin, Jessica Colnago, Lorrie Faith Cranor et al. "Design and evaluation of a data-driven password meter." In Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, pp. 3775-3786. ACM, 2017.

[49] Ur, Blase, Fumiko Noma, Jonathan Bees, Sean M. Segreti, Richard Shay, Lujo Bauer, Nicolas Christin, and Lorrie Faith Cranor. "" I Added'!'at the End to Make

It Secure": Observing Password Creation in the Lab." In Eleventh Symposium On Usable Privacy and Security ({SOUPS} 2015), pp. 123-140. 2015.

[50] Ur, Blase, Jonathan Bees, Sean M. Segreti, Lujo Bauer, Nicolas Christin, and Lorrie Faith Cranor. "Do users' perceptions of password security match reality?." In Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, pp. 3748-3760. ACM, 2016.

[51] Ur, Blase, Patrick Gage Kelley, Saranga Komanduri, Joel Lee, Michael Maass, Michelle L. Mazurek, Timothy Passaro et al. "How does your password measure up? the effect of strength meters on password creation." In Presented as part of the 21st {USENIX} Security Symposium ({USENIX} Security 12), pp. 65-80. 2012.

[52] Verizon Enterprise Solutions. 2016. 2016 Data Breach Investigations Report Retrieved February 19, 2020 from https://enterprise.verizon.com/resources/reports/2016/dbir-2016-executive-summary.pdf

[53] Virvou, Maria, George Katsionis, and Konstantinos Manos. "Combining software games with education: Evaluation of its educational effectiveness." Journal of Educational Technology & Society 8.2 (2005).

[54] Vu, Kim-Phuong L., Robert W. Proctor, Abhilasha Bhargav-Spantzel, Bik-Lam Belin Tai, Joshua Cook, and E. Eugene Schultz. "Improving password security and memorability to protect personal and organizational information." International Journal of Human-Computer Studies 65, no. 8 (2007): 744-757.

[55] Weirich, Dirk, and Martina Angela Sasse. "Pretty good persuasion: a first step towards effective password security in the real world." Proceedings of the 2001 workshop on New security paradigms. ACM.

[56] Yan, Jianxin Jeff. "A note on proactive password checking." Proceedings of the 2001 workshop on New security paradigms. ACM, 2001.

[57] Zviran, Moshe, and William J. Haga. "Password security: an empirical study." Journal of Management Information Systems15.4 (1999): 161-185.

[58] Hunt, T. (2019). Have I Been Pwned. Retrieved from https://haveibeenpwned.com/Passwords

Appendices

A. Demographics Survey

We conducted an initial demographic survey of our participants. The questions asked are as follows:

a. Do you have background in CS/IT/Security?

a. Yes b. No

b. **What is the highest level of education you have completed?**

a. Bachelor's Degree

b. Master's Degree

c. Doctorate

d. Other

c. **What is your Gender**

a. Female

b. Male

c. Other

d. **How old are you?**

a. 21-30

b. 31-40

c. 41-50

d. Above 50

e. **Select your Nationality** (from the list of nations).

B. Game Tests

The questions asked during the pre- and post-tests were focused on the heuristics being taught in the game. Every question contained two password choices, with one satisfying the particular heuristic, while the other failed to do so. We created the password choices by picking suitable passwords from leaked databases [1, 58] and minimally modifying them to be able to test a particular heuristic (similar to the method followed in [50]). For example, we chose a leaked password "*Password*", (leaked 130,999 times as per [58]), and incorporated repeated characters (H6) by adding four consecutive "s" to create "*Passssword*". This password did not satisfy H6. The corresponding alternate choice for this password was

"*Paswwoorsdd*", which had a maximum of two repeated characters. The latter satisfied the requirements for the heuristic H6. In this pair, the former is considered the weaker password. Similarly, we chose other leaked passwords and incorporated the same methodology to create a list of password pairs for use in the pre- and post-tests.

The passwords given as choices were judged based on the particular heuristic being dominant. For example, considering passwords "*asdfghjkl*" and "*afhsgdljk*", despite having similar characteristics, the former is a very common **keyboard pattern** and the latter is not. This puts the former in "common passwords" criteria, making it weaker. Both the options provided were of equal length and had similar characteristics like order of character classes etc., except the heuristic being checked. Presence of multiple heuristics within a password would have caused ambiguity to the participants. Apart from judging based on heuristics, we also analyzed the password strengths using several available tools to ensure that our choice of answers are true to the maximum extend. Following online services were used to rate the passwords before being added to the pre and post-tests:

My1Login.2020. Take the Password Test. Retrieved June 02, 2020 from https://www.my1login.com/resources/password-strength-test/

Kaspersky.2020.Kaspersky Secure Password Check. Retrieved June 02, 2020 from https://pasword.kaspersky.com/in/

The level-wise test questions and the heuristic being checked (*in italics*) are provided below:

Which of the following passwords is weaker?

Level 1 Pre-Test

1. a) **Passsssword** *(Repeated characters, H6)*
   b) Paswwoorsdd
   c) Both are identical

2. a) welccoommlee
   b) **weeleecomeee** *(Duplicate Character, H7)*
   c) Both are identical

3. a) **Passtuvw@12rd** *(Alphabetic Sequence, H9)*
   b) Pstasuwv@12rd
   c) Both are identical

4. a) ac2ab1c12bacbca
   b) **abc12abc12abcabc** *(Repeated Section, H8)*
   c) Both are identical

Level 1 Post-Test

1. a) **seeseameee** *(Duplicate Character, H7)*
   b) seasamemesa
   c) Both are identical

2. a) **tcstcstcsadmintcs** *(Repeated Section, H8)*
   b) sctatdmcicnstntcs
   c) Both are identical

3. a) letmmelein
   b) **letmmmmmin** *(Repeated Character, H6)*
   c) Both are identical

4. a) **Nutriopqrst** *(Alphabetic Sequence, H9)*
   b) Nutriposqtr
   c) Both are identical

Level 2 Pre-Test

1. a) jo14n21ny
   b) **jonny1421** *(Digit Predictable, H12)*
   c) Both are identical

2. a) **Brooklyn** *(Uppercase predictable, H13)*
   b) brooklYn
   c) Both are identical

3. a) **RockWell@789** *(Symbol Predictable, H11)*
   b) R@ockwell789
   c) Both are identical

4. a) tomhjklinmpo

   b) **tomhijklmnop** *(Alphabetic Sequence, H9)*

   c) Both are identical

5. a) **qwertyuiop** *(Keyboard Pattern, H10)*

   b) qtrwyeuiop

   c) Both are identical

6. a) **pass@July2017** *(Date Format, H15)*

   b) pass@Ju1y2@17

   c) Both are identical

Level 2 Post-Test

1. a) **asdfghjkl** *(Keyboard Pattern, H10)*

   b) afhsgdljk

   c) Both are identical

2. a) shad7w00o

   b) **shadow007** *(Digit Predictable, H12)*

   c) Both are identical

3. a) **Me@Aug95** *(Date Format, H15)*

   b) Me@Au9g5

   c) Both are identical

4. a) **secure@234** *(Symbol Predictable, H11)*

   b) sec@23ure4

   c) Both are identical

5. a) **Can@da** *(Uppercase predictable, H13)*

   b) caN@da

   c) Both are identical

6. a) sutjerrywv

   b) **stuvwjerry** *(Alphabetic Sequence, H9)*

   c) Both are identical

For each of these test questions, there was another question asking the player's confidence level for their responses:

How confident are you about your selection:

a. Not at all confident

b. Not very confident

c. Neutral

d. Confident

e. Very confident

C. Game Feedback Survey

Please answer all these questions:

1) The game was fun

   a. Strongly agree

   b. Agree

   c. Neutral

   d. Disagree

   e. Strongly Disagree

2) The game was educational

   a. Strongly agree

   b. Agree

   c. Neutral

   d. Disagree

   e. Strongly Disagree

3) I learned how to create a secure password

   a. Strongly agree

   b. Agree

   c. Neutral

   d. Disagree

   e. Strongly Disagree

D. Information Regarding our Data and RockYou database:

In the game, the participants have created passwords that satisfy all the game heuristics in Level 2. These can be considered as passwords with lesser vulnerabilities.

Table A shows the list of top 10 structures from Level 2 and its corresponding occurrence in the RockYou database of over 32 million leaked passwords [1].

RockYou database consists of leaked passwords from various sources. The password structures that cleared Level 2 were not found in the leaked database. The most common passwords from RockYou database included structures with very little heuristics incorporated, like LLLLLL (12.23%), LLLLLLL (8.35%), LLLLLLLL (7.57%), DDDDDD (6.98%) etc. This shows that the leaked passwords possibly satisfied fewer heuristics than Level 2 passwords.

Password Structures (Level 2)	Occurrences (%)	
	Game Data	RockYou Data
LSLLLDLU	0.300	0.00003
LSLLLDLLU	0.115	0
LLLUSDLLL	0.057	0.000003
ULLSULDDLL	0.046	0
LLLUSDDDUU	0.046	0
UUUUSDDLL	0.035	0
USLSUDLU	0.035	0
SLLDLLLU	0.035	0.000003
LULLLSULLDDDL	0.029	0
ULSULDDDDL	0.029	0

Table A: Top 10 recurring password structures from the study vs. RockYou data

E. Animals in Game used to Teach and Check Password Heuristics

The game teaches password heuristics using oncoming animals, who provide heuristic-related information to the players, at the same time, check if heuristics are satisfied during password creation. Given below are the various animals that come into play. Their corresponding heuristics are mentioned in Table 1.

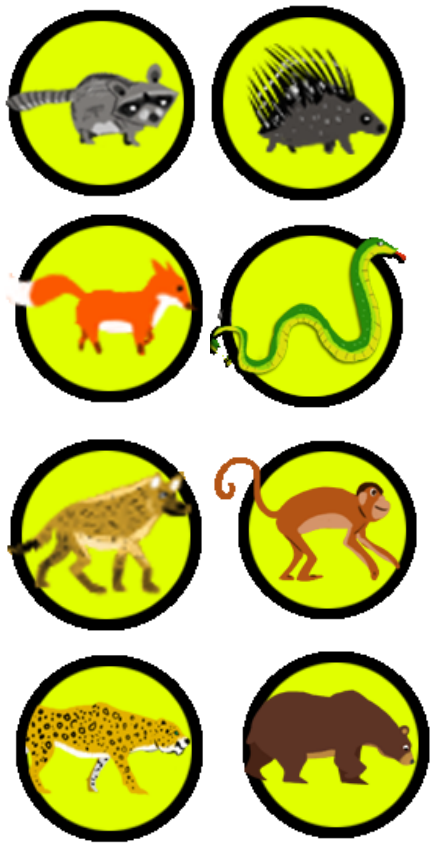

Figure A: The animals appearing in the game are (starting from top left) Raccoon, Porcupine, Fox, Snake, Hyena, Monkey, Leopard, and Bear.