An evaluation of the Game Changer Password System: A new approach to password security

Conor T. M^{c}Lennan[a,*], Philip Manning[b], and Samantha E. Tuft[a]

[a]Department of Psychology, Language Research Laboratory, Cleveland State Univiersity, 2121 Euclid Ave., Cleveland, OH 44115, USA

[b]Department of Criminology, Anthropology, and Sociology, Cleveland State Univiersity, 2121 Euclid Ave., Cleveland, OH 44115, USA

Email addresses: c.mclennan@csuohio.edu (C.T. M^{c}Lennan), p.manning@csuohio.edu (P. Manning), s.tuft@vikes.csuohio.edu (S.E. Tuft).

*Corresponding author

Abstract

We propose – and experimentally test – a mnemonic variant of password security that uses game positions as passwords. In Experiment 1, we report accuracy and reaction time data when high school student, younger adult, and older adult participants remembered and entered one game-based password, using chess or Monopoly. In Experiment 2, we report accuracy and reaction time data from participants' use of five game-based passwords across 24 sessions over 10 weeks. All five passwords were stored in chess or Monopoly for the initial 20 sessions, and changed (from chess to Monopoly or vice versa) for the remaining sessions. This new approach to password security is both mathematically robust and user-friendly.

Keywords: password security; memory; interference; cognitive

1. Introduction

It has become commonplace to note that computer systems protected by authentication systems are either mathematically robust or user-friendly, but not both. As Taneski, Hericko, and Brumen (2014) recently reported, "The computer community has not made a very much-needed shift in password management for more than 35 years" (page 1360). Indeed, it is time for a new approach. We propose a mnemonic variant of password security in which users place game pieces at various positions on a game board in order to create game-based passwords. In doing so, we are building off of previous research that used games for passwords (see Tao & Adams, 2008; Malempati & Mogalla, 2011; Pandey, 2012). This new approach to password security (1) enhances the mathematical advantages of other systems, (2) can be presented as a *fun* solution to cybersecurity threats, and can therefore be user-friendly, and (3) can facilitate users' willingness to change passwords as their new choices are unlikely to interfere with memories of their previous ones. In our view, this cognitive-based approach to cybersecurity is an important new tool to facilitate user compliance. We acknowledge that future cybersecurity systems will involve multiple authentication checks, as recommended by many researchers (for example, Lo, 2016). Our contribution is therefore to describe one piece in the jigsaw puzzle that safeguards gateways to both the Internet and other electronic highways.

Consider the following scenario: You sit down to balance your checkbook, go to your bank's website, and rather than typing in an alphanumeric password, you see a panel of games. From this panel, you select chess, a chessboard pops up (e.g., see the Appendix), and you move the black king to the bottom left square and the white queen three positions above where you placed the black king, and voilà, you are in. Our approach is envisioned as one in which a number of different games are presented on the screen, similar to viewing a screen full of movie options in Netflix, allowing users to select the game first - and then to enter his or her password in the game selected. Doing so will only require one extra step for the user (remember which game you have stored your password) while significantly increasing the underlying security. Of course, the number of different games from which to choose would depend on a number of different factors, including level of security needed and user preference. The current study provides a starting point for such a framework. That is, prior to building such a platform, we conducted experiments designed to examine participants' ability to remember, enter, and use game-based passwords.

1.1 Overview

We begin by briefly discussing the main contributions the current paper makes to the literature on password security by paying particular attention to the use of password security in computers from a psychological perspective. We then introduce the Game Changer Password System, a new approach to password security motivated in large part from basic research in cognitive psychology, as well as other work using graphical passwords.

In Sections 2 and 3, we describe the experimental testing of this new password system. In Section 4, we discuss the Game Changer Password System based on the reported empirical data. We then describe some limitations and future work. We finish in Section 5 by providing several concluding remarks.

1.2 Contributions to Password Security Literature

In the current paper, we make the following contributions to the literature on password security:

1) We present the Game Changer Password System, a novel approach to password security in which the game-based passwords are both highly secure and easy to remember.
2) In Experiment 1, we report data from 131 participants' use of two game-based passwords, chess and Monopoly, across three different age groups during a single session.
3) In Experiment 2, we report data from 33 participants' use of multiple game-based passwords across 24 sessions over the course of 10 weeks.

1.3 The Game Changer Password System

The approach to password protection proposed in this paper stores passwords in game positions[1]. This approach involves giving up the idea of passwords as alphanumeric strings and replacing such strings with iconic codes that are stored on the virtual game positions of different games.

1.3.1 Mathematical Robustness and Security

The use of game boards promises to enhance cybersecurity because game boards offer an array of advantages over traditional approaches to password protection. Mathematically, game boards store passwords multi-dimensionally – because each item of the password has to be located on a particular square or location. The transformation is astounding: a four digit numeric code (as used on an iPhone) has 10,000 combinations. By contrast, a four-icon code stored on a chessboard has more than 316 billion combinations (explained further below). Moreover, with enhancements and the availability of multiple games, the combinations increase dramatically.

[1] The Game Changer Password System was described in U.S. Provisional Patent No. 61782062, 2013, and in United States National Science Foundation Grant No: 1343141.

Let us consider the combinations more closely. For our current experimental testing and implementation of the Game Changer Password System, we chose chess and Monopoly. For chess, there were 12 pieces and 64 squares (or locations) in which the participants could choose to create their passwords. For Monopoly, there were 7 pieces[2] and 40 locations in which participants could choose to create their passwords. Participants were asked to create a password using either two or four game pieces on the board. Given that pieces could be reused, but locations could only be used once, the combinations for chess with passwords consisting of four and two game pieces are 316,203,761,664 and 580,608, respectively. For Monopoly, the combinations with passwords consisting of four and two game pieces are 5,266,257,360 and 76,440, respectively. As in chess, pieces could be reused, but locations could only be used once. The decrement in performance between the two and four piece conditions may be nowhere near the increment in security gained. Similarly, gains can be made with more pieces, locations, or both to choose from. Imagine having four colors of chess pieces instead of two. The number of combinations for a two-piece password would increase from 580,608 to 2,322,432 – an increase in security while still only requiring the user to put two pieces on the board.

The Game Changer Password System is mathematically appealing because it is multi-dimensional and hence capable of generating large reservoirs of possible passwords from simple

[2] There were actually 8 game pieces shown, which would significantly increase the number of combinations compared to 7 game pieces. However, unfortunately the shoe game piece was shown twice, and this error was not detected before data collection began. Consequently, calculations are based on 7 pieces. Also, please keep this caveat in mind when examining the pieces count shown in the Appendix C.

starting points, such as a chess or Monopoly board. Nevertheless, this password system – like any password system – will be vulnerable to unthrottled guessing when up against a powerful computer that can race through millions of combinations. For example, automated brute force attacks may take only minutes (see Brumen & Cernezel, 2014; Brumen & Taneski, 2015; van Oorschot, Salehi-Abari, & Thorpe, 2010). However, as Gao, Jia, Ye, and Ma (2013) report, "Generally speaking, it is more difficult to use a brute force attack against graphical passwords than alphabetic schemes" (page 1689). Moreover, with appropriate additional, and easily implemented, safeguards, the multi-dimensionality of game approaches is robust.

There are two additional points related to this discussion of increased security, *stacking* and *layering*. Imagine, in Monopoly, if users were allowed to put more than one piece on a given square. This notion of stacking makes sense in Monopoly, but not chess – but only based on the rules of the games. Recall that a major advantage of the Game Changer Password System is that users are not required to know how to play the game – or follow the rules even if they do know how to play. So, stacking could work for chess (or any other game) as well. The four game piece password in Monopoly with stacking would increase from 5,266,257,360 to 6,146,560,000, just by allowing users to reuse – or place more than one game piece – in any given location.

Layering, which provides an even greater contribution to the mathematical robustness of the system, refers to the possibility of allowing (or requiring) users to place pieces on two different game boards as their password. So, for example, he or she would first place two pieces on a Monopoly board and then place two pieces on a chessboard. The resulting number of combinations is 76,440 * 580,608 = 44,381,675,520. If the user first had to select Monopoly out of a screen with 10 different games and subsequently select chess, in line with the Netflix analogy described earlier, security would increase further (especially if the order of the two or more games mattered). Of course, the amount of layering, the decision to allow stacking or not, the number of games to choose from, whether or not the order of games selected should matter all depend on the amount of security desired. Moreover, these decisions are not mutually exclusive. The point here is that the number of combinations (or permutations, if order mattered) could quickly become astronomical for a password that participants should still be able to remember relatively easily and enter relatively quickly.

Future investigations will need to evaluate the impact that stacking, layering, or both has on usability. However, even if a greater number of pieces, stacking, layering, or some combination is required in order to increase the brute-force resilience of the system, we believe that the advantages of the Game Changer Password System outweigh any possible problems, such as a potential reduction in memorability. In particular, two such advantages are that the system is fun and usable by different populations of users, both of which are described in greater detail later in the paper.

The calculations underlying the mathematical robustness are based on the complete theoretical password space. It is, of course, important to consider the effective password space; that is, passwords users are most likely to choose (Gao, Jia, Ye, & Ma, 2013). Future calculations for game-based passwords could be ongoing based on effective password space. Moreover, just as researchers have developed effective password strength meters for alphanumeric passwords (Wang, He, Cheng, & Wang, 2016), future efforts could be devoted to developing such meters for game-based passwords. Future investigations should consider directly comparing the Game Based Password System with traditional alphanumeric or textual passwords. For now, it is worth noting that graphical password systems typically provide a theoretical password space similar to

– or larger than – the theoretical password space of alphanumeric systems (Suo, Zhu, & Owen, 2005). Direct comparisons of the password spaces offered by the Game Based Password System and other graphical password systems are complicated for three reasons. First, there is considerable variability in the password spaces of existing graphical systems. Second, since this is the first investigation of the Game Based Password System, it is premature to calculate the effective password space. Finally, both the theoretical and effective password spaces of the Game Based Password System would change significantly depending on a number of decision points, such as the number of pieces required, in addition to whether stacking, layering, or both were implemented.

1.3.2 Comparisons with Other Game-Based Systems

Although there is a voluminous literature on graphical passwords (e.g., see Gao, Jia, Ye, & Ma, 2013; Gao, Ma, Jia, & Ye, 2012; Zakaria, Griffiths, Brostoff, & Yan, 2011; Ramapriya, Yamini, Moorthy, & Scholar, 2015), to our knowledge, there are only three other research groups that have attempted to use games to secure passwords. Tao and Adams (2008) used the Chinese game of Go; Malempati and Mogalla (2011) conducted a relatively small user study with Snakes and Ladders; and Pandey (2012) adapted chess by focusing on one of two pieces – either the rook or the bishop. Tao and Adams (2008) report 78% accuracy over 13 weeks and 65% accuracy rate for the first week. Malempati and Mogalla (2011) reported that Snakes and Ladders was promising, a conclusion also reached by Pandey (2012) for chess. These are valuable precursors to the approach proposed in the current paper because they demonstrate the viability of the use of game positions as passwords. However, these earlier papers limited their approach to a particular game and therefore did not explore the advantages of making multiple games available to users, nor did they measure login time or report any statistical analyses.

Many of the previous systems were not evaluated to the extent that we are doing, as other security researchers have called for (e.g., Biddle, Chiasson, & van Oorschot, 2012), and thus lack the appropriate empirical data for direct comparisons. Furthermore, our approach does not require participants to learn rules or have any knowledge of how to play the games. Indeed, users could even design and create their own games. Moreover, the crucial game changing aspect of our approach is that users select different games a.) over time (e.g., when they are required, or decide, to change their password), b.) for different devices, c.) for different accounts, or d.) for any combination of a.) – c.). This last feature is crucial because changing the game makes use of different types of password, with comparable security, and gives users some choice regarding their authentication method, as advocated by Cheswick (2013).

1.3.3 Advantages of the Game Changer Password System

In addition to multi-dimensionality, the use of games is able to address a variety of memory-related problems in a way that no existing password scheme is able to do. First and most straightforwardly, visual game-based passwords should be easier to remember than alphanumeric ones, given the well-documented Picture Superiority Effect in the memory literature. The Picture Superiority Effect refers to the finding that memory is typically better for pictures than other types of information, including words (e.g., Paivio, & Csapo, 1973; Khan, Aalsalem, & Xiang, 2011; see however, De Angeli, Coventry, Johnson, & Renaud, 2005). Secondly – and ironically – multiple iconic passwords are easier to separate from each other. As mentioned above, the use of multiple passwords could occur because a user is required to change an existing password, has different passwords for different devices, has different passwords for different accounts, or any combination of these reasons. As a routine security measure, users are often required periodically to select new passwords (see Wang & Wang, 2016). This requirement is often met with user resistance and frustration. As a result, the users' *new* passwords are often quite close to their *old* ones (Wang, He, Cheng, & Wang, 2016). Such similarity is bad both for security and because memory interference ensues, as users mix up their different – but similar – passwords, not to mention security concerns associated with choosing passwords that correlate with usernames, emails, and other personal information (Ji, Yang, Hu, Han, Li, & Beyah, 2016).

Psychologists have long known that people may have trouble retrieving information from long-term memory because of interference with similar memories (Underwood, 1957), and this notion of interference has been described and documented in cognitive psychology (e.g., in spoken word recognition, Vitevitch & Luce, 1998) and neuroscience (Engelmann, 2009). Moreover, the role that interference may play in users' ability to remember their passwords is becoming increasingly well known to password security researchers. For example, Biddle et al. (2012) notes the following. "For usability, a major concern is multiple password interference" and "Looking ahead, we expect that tomorrow's ideal graphical password systems may have … design features minimizing password interference" (pages 36-37). Indeed, a game-based approach may address this issue by allowing users to store new passwords in new games, as in the example below, thereby reducing the risk of interference.

Consider two games as examples and test cases: chess and Monopoly. To set an iconic code in the first case, the user puts a number of pieces – say four – on a virtual chessboard. His or her password could be, then, four white pawns in a row. This password could be entered by a mouse or just moved by a finger dragging the icon across the screen. Then imagine letting a month – or six – go by. The system – a bank account or perhaps a workplace protocol – requires the user to select a new password. Flipping through a virtual catalog of games, the user selects

Monopoly and proceeds to put three hotels on Park Avenue and a hat in jail. More creatively, perhaps, a user could put three dogs on GO and turn the Community Chest card to *You have won third prize in a beauty contest! Collect $30!* One of these Monopoly *iconic codes* becomes the new password until the system requires another change, or until the user opts to do so.

For our current experimental testing and implementation of the Game Changer Password System, we chose chess and Monopoly. These games were selected to minimize the similarity between the games. Chess and backgammon, for example, may be more likely to lead to interference, and thus would not benefit as much from a game change from one to the other. There is no limit to the number of games that could be used. Indeed, as mentioned previously, the approach does not require a real or existing game. All that is required is a grid with pieces or icons. So, for example, we could imagine a grid of different fish tanks with some number of different fish icons that would be selected and placed in the chosen tanks.

Many security researchers have assumed a necessary tradeoff between security and usability (e.g., memorability), such that increasing one decreases the other (Bonneau, Herley, van Oorschot, & Stajano, 2012; 2015). As some prominent researchers put it: "…it is important to determine methods for password generation that will yield passwords that provide adequate security but are also memorable" (Vu, Proctor, Bhargav-Spantzel, Tai, Cook, et al., 2007, page 745) and "...graphical password systems provide an alternative ... to alphanumeric strings that should be explored more fully because they exploit people's ability to recognize pictures …with high accuracy..." (Vu, et al., 2007, page 756). We agree with Herley and van Oorschot (2012) that "…conventional security wisdom oversimplifies the story to a tradeoff between security and usability." We further agree with Herley and van Oorschot (2012) that finding a silver bullet that addresses all relevant password issues, including security and usability, is unlikely. Nevertheless, Tari, Ozok, and Holden (2006) found that graphical passwords could lead to less vulnerability than traditional alphanumeric passwords, and Rao, Pravallika, Priyanka, and Kumar (2016) recently proposed a shoulder-surfing resistant graphical password scheme. Furthermore, Chiasson, Forget, Stobert, van Oorschot, and Biddle (2009) found click-based graphical passwords were significantly less susceptible to multiple password interference than text passwords. Widenbeck, Waters, Birget, Brodskiy, and Memon (2005) found that while practicing entering their passwords, the graphical users took longer and were less accurate at inputting their passwords than the alphanumeric users. Additionally, the two groups did not differ in their memory of their password over weeks without use, but the graphical users spent more time entering a password than the alphanumeric users. Therefore, we suspect — and present data to support the claim — that any decrement in performance (e.g., slightly longer times to enter a game-based password) is more than compensated by the increment in security gained from using the Game Changer Password System. However, since this is an empirical question, we conducted experiments to obtain objective data regarding users' memory for, and use of, this novel password scheme.

Bonneau et al. (2012) also highlight the importance of understanding that cybersecurity is as much a human accomplishment as it is as a technological breakthrough. End-users have to be willing and able participants in the protection of their own information. As will be discussed later, one essential condition for the acquisition of compliance is usability. For example, a string of 30 unrelated keyboard symbols might be robust but it is not user-friendly. The point here is that it is important to keep the user in mind (Whalen, 2011). Recent approaches attempt to do so by helping users develop effective password creation schemes (Curran & Snodgrass, 2015), by providing feedback on passwords that users create (Shay, et al., 2015), by developing more robust password storage systems (Eich, John, Smith, & Cankaya, 2016), or by investigating factors that encourage users to adopt password managers (Alkaldi & Renaud, 2016). Another recent study found that providing guidance on how to avoid insecure PINs resulted in users choosing more secure PIN choices (Gutmann, Volkamer, & Renaud, 2016).

However, there are two important factors that are overlooked by Bonneau et al. (2012). The first is that user compliance can be achieved by *acquiescence through enjoyment*. That is, by ensuring that users have some fun while interfacing with devices that require end-user authentication, users are more likely to be compliant and more accurate in their use of cybersecurity measures. As Bonneau, Herley, van Oorschot, and Stajano (2015) point out, “Users are less likely to buy in to any system that presents them with inconveniences they do not understand” (page 84). In our view, one key way to increase users’ compliance is to offer users a system that entertains them a little. Games, we propose, offer this possibility – albeit with some restrictions. Although it may be hard to imagine how any password system can be fun, that is where games, literally and metaphorically, come into play. As the following quote illustrates, we are not alone in wanting a password system that can be fun. “My dream is that authentication might become a lot less odious, maybe even fun” (Cheswick, 2013, p 4).

The second factor overlooked in Bonneau et al. (2012) is the *relative usability* of password schemes in different populations. The phrase *relative usability* is a helpful reminder that what one group finds child’s play another sees as rocket science. A possible strength of the Game Changer Password System advocated here is that it may be appealing to older populations who are often late adopters of new technologies, and are often a neglected population in password security research.

In the next section we report the experimental testing of the Game Changer Password

System. In particular, we examined participants' memory for newly created passwords in chess and Monopoly, across three different age groups: high school students, younger adults, and older adults. Our focus on older adults represents an important shift. Vu and Hills (2013) suggest that pictures could be used as cues for recalling passwords for older adults. They found that older adults were more likely to recall passwords that were generated using an image-based mnemonic technique compared to a text-based mnemonic technique. Older adults are often a neglected population in password security research, especially in graphical password systems, which is surprising given the aging population. According to the US Department of Health and Human Services Administration on Aging (2010), people 65 and over represented 12.4% of the U.S. population in 2000 and are expected to grow to 19% by 2030. We also report how long it takes participants to enter their passwords. Intuitively, it would be less pleasant to place game pieces on a game board using a smaller device or screen size. However, we also wanted to know whether a device with a smaller screen in which the game boards were intentionally too large for the display, and thus had to be moved around in order to see, and use, the entire board and the game pieces, would also lead to performance costs. More specifically, having to do so would presumably take longer, but whether or not participants would actually make more errors was one of the empirical questions addressed in this experiment.

2. Experiment 1

2.1 Method

2.1.1 Participants

One hundred and thirty-one participants were recruited from the Cleveland area. Forty-three high school students (21 females) between 14 and 16 years of age[3] were recruited from two local high schools. In addition, 48 younger adults (32 females) between 18 and 34 years of age and 40 older adults (20 females) between 60 and 85 years of age were recruited from the Cleveland State University community. Participants either received partial research participation credit or monetary compensation for their participation per our approved Institutional Review Board protocol.

2.1.2 Materials and Procedure

Participants were first asked to complete a consent form. Participants were then given either an iPod Touch (the same size and shape as an iPhone) or an iPad Mini with either a chessboard or a Monopoly board application open. Examples of the chessboard and Monopoly boards are shown in Appendices B and C.

Participants were asked to create a password that they thought was secure and that they would use as an actual password by placing either two or four game pieces on the board. After the participants created their passwords, there was a 10-20 minute delay (cognitive psychologists consider anything beyond approximately 20 seconds long-term memory), during which

[3] One participant’s age and gender were not reported.

participants were asked to complete a questionnaire about games and passwords[4] and additional filler tasks (in order to fill the delay and minimize the likelihood that participants were rehearsing their newly-created passwords). After the delay, participants were asked to re-enter their password (and the correct combination of pieces and locations was required). Participants were not specifically told in advance that they would be asked to do so. Participants had up to three chances to correctly enter their password. Reaction time (RT) and percentage correct (PC) were recorded for each participant. RTs were measured in seconds (s) from the onset of participants' initial game piece movement until the *Check Password* button was clicked. RTs are only meaningful, and thus are only reported and analyzed, for correct entries.

2.1.3 Design

Password Size (two or four pieces), Game (Chess or Monopoly), and Device (iPod Touch or iPad Mini) were all completely counterbalanced in all three groups of participants (High school Students, Younger Adults, and Older Adults). Thus, including Age Group, we used a 3 × 2 × 2 × 2 completely between participants design.

2.2 Results

Two separate 3 (Age Group: High School Students, Younger Adults, Older Adults) × 2 (Password Size: two, four) × 2 (Game: Chess, Monopoly) × 2 (Device: iPod Touch, iPad Mini) completely between participants ANOVAs were performed, one on PCs within three attempts and one on RTs to correct attempts. The mean percentages correct (PCs) and the mean reaction times (RTs) as a function of Age Group, Password Size, Game, and Device within three attempts are reported in Tables 1 and 2, respectively.

[4] See Appendix A for the questions that were asked.

2.2.1 Accuracy

101 of the 131 participants (77%) entered their passwords correctly within three attempts. 85 of those 101 participants (83%) entered their passwords correctly on the first attempt. Fourteen of those 101 participants (14%) entered their passwords correctly on the second attempt. Three of those 101 participants (3%) entered their passwords correctly on the third attempt.

The main effect of Age Group was not significant, $F(2, 107) = 2.49$, $MSE = 1747.66$, $p = .088$, $\eta_p^2 = .04$. Using the Bonferroni correction, there was no difference in accurately entering passwords within three attempts between high school students (M = 69%, SE = 6%), younger adults (M = 88%, SE = 6%), and older adults (M = 73%, SE = 7%).

The main effect of Game was not significant, $F(1, 107) < 1$, $MSE = 1747.66$, $p = .50$, $\eta_p^2 < .01$. The mean PCs for Chess and Monopoly were 74% and 79%, respectively.

The main effect of Password Size was significant, $F(1, 107) = 7.41$, $MSE = 1747.66$, $p = .008$, $\eta_p^2 = .07$. More specifically, participants with a password consisting of two pieces (*M* = 86%, *SE* = 5%) were more accurate than participants with passwords consisting of four pieces (*M* = 66%, *SE* = 5%).

Table 1

Mean Percentages Correct (PCs) and Standard Errors (SEs) as a Function of Age Group, Password Size, Game, and Device Within Three Attempts

			Game					
			Chess		***Monopoly***		***Overall***	
Age Group	**Device**	**Password Size**	**M**	**SE**	**M**	**SE**	**M**	**SE**
High School Students	iPod Touch	Two pieces	50	17	83	17	67	12
		Four pieces	40	19	50	21	45	14
	iPad Mini	Two pieces	100	19	80	19	90	13
		Four pieces	67	17	83	19	75	12
Younger Adults	iPod Touch	Two pieces	100	17	100	17	100	12
		Four pieces	83	17	50	17	67	12
	iPad Mini	Two pieces	83	17	100	17	92	12
		Four pieces	83	17	100	17	92	12
Older Adults	iPod Touch	Two pieces	80	19	100	19	90	13
		Four pieces	60	19	60	19	60	13
	iPad Mini	Two pieces	80	19	80	19	80	13
		Four pieces	60	19	60	19	60	13
		Overall	74	5	79	5		

The main effect of Device was not significant, *F*(1, 107) = 1.85, *MSE* = 1747.66, *p* = .18, η_p^2 = .02, although participants' mean PC to passwords on the iPad Mini was 81%, which was a trend of 10% better performance than on the iPod Touch (71%). No other effects of PC were significant (all *p*s > .19).

Before moving on to an analysis of the reaction time data, we first consider the types of errors that were made. In particular, separately for each age group, we report

the percentage of errors for the following four different types of errors: 1.) A location only error occurred when not all of the locations were selected correctly, but all of the correct pieces were selected. 2.) A piece only error occurred when not all of the pieces were selected correctly, but all of the correct locations were selected. 3.) A location and piece error occurred when not all of the correct locations were selected and not all of the pieces were selected correctly. 4.) Finally, what we refer to as a combination error occurred when all of the correct pieces and all of the correct locations were used but some combination of pieces and locations was incorrect.

For high school students, 7.14% of the errors involved location only, 33.33% of the errors involved piece only, 50.00% involved both location and piece, and 9.52% involved a combination. For younger adults, 30.43% of the errors involved location only, 34.78% of the errors involved piece only, 21.74% involved both location and piece, and 13.04% involved a combination. Finally, for older adults, 35.71% of the errors involved location only, 26.19% of the errors involved piece only, 21.43% involved both location and piece, and 16.67% involved a combination.

2.2.2 Reaction Time

The overall mean RT to correctly enter a password was 28s (*SE* = 3s). The main effect of Age Group was significant, $F(2, 77) = 11.53$, $MSE = 826.63$, $p < .001$, $\eta_p^2 = .23$. Using the Bonferroni correction, high school students ($M = 17$s, $SE = 6$s) were significantly faster at entering their password correctly than older adults ($M = 49$s, $SE = 5$s), $p < .001$, but not younger adults ($M = 19$s, $SE = 5$s), $p > .99$. Younger adults were significantly faster at entering their password correctly than older adults, $p < .001$.

The main effect of Game was marginally significant, $F(1, 77) = 3.65$, $MSE = 826.63$, $p = .06$, $\eta_p^2 = .05$. Specifically, there was a trend for Chess passwords ($M = 34$s, $SE = 4$s) to be correctly entered more slowly than Monopoly passwords ($M = 22$s, $SE = 4$s).

The main effect of Password Size was significant, $F(1, 77) = 6.41$, $MSE = 826.63$, $p = .013$, $\eta_p^2 = .08$. More specifically, participants with a password consisting of two pieces (*M* =

20s, *SE* = 4s) entered their password correctly significantly faster than participants with passwords consisting of four pieces (*M* = 36s, *SE* = 5s).

The main effect of Device was significant, *F*(1, 77) = 16.54, *MSE* = 826.63, $p < .001$, $\eta_p^2 = .18$. More specifically, participants' who correctly entered their password using an iPod Touch (*M* = 40s, *SE* = 4s) were significantly slower than participants' using an iPad Mini (*M* = 16s, *SE* = 4s).

The Age Group × Game interaction was significant, *F*(2, 77) = 3.15, *MSE* = 826.63, *p* = .048, $\eta_p^2 = .08$. Simple effects for Game were preformed using the Bonferroni correction. For correctly entered Chess passwords, high school students (*M* = 17s, *SE* = 8s) were not significantly faster than younger adults (*M* = 19s, *SE* = 6s), $p > .99$. However, both high school students and younger adults were significantly faster at correctly entering their Chess passwords than older adults (*M* = 65s, *SE* = 8s), *p*s < .001. For correctly entered Monopoly passwords, there were no significant differences in mean RTs between high school students (*M* = 17s, *SE* = 8s), younger adults (*M* = 18, *SE* = 7s), and older adults (*M* = 32s, *SE* = 8s), *p*s > .45.

Table 2

Mean Reaction Times (RTs) in Seconds (s) and Standard Errors (SEs) as a Function of Age Group, Password Size, Game, and Device Within Three Attempts

			Game					
			Chess		***Monopoly***		***Overall***	
Age Group	**Device**	**Password Size**	**M**	**SE**	**M**	**SE**	**M**	**SE**
High School Students	iPod Touch	Two pieces	16	17	21	13	18	10
		Four pieces	27	20	26	20	27	14
	iPad Mini	Two pieces	10	13	7	14	8	10
		Four pieces	14	14	12	13	13	10
Young Adults	iPod Touch	Two pieces	25	12	14	12	19	8
		Four pieces	27	13	28	17	28	11
	iPad Mini	Two pieces	8	13	12	12	10	9
		Four pieces	17	13	18	12	18	9
Older Adults	iPod Touch	Two pieces	56	14	36	13	46	10
		Four pieces	145	17	63	17	104	12
	iPad Mini	Two pieces	30	14	10	14	20	10
		Four pieces	31	17	20	17	25	12
		Overall	34	4	22	4		

The Age Group × Device interaction was significant, $F(2, 77) = 5.19$, $MSE = 826.63$, $p = .008$, $\eta_p^2 = .12$. Simple effects for Device were preformed using the Bonferroni correction. For correctly entered passwords using the iPod Touch, high school students (M = 22s, SE = 9s) were not significantly faster than younger adults (M = 23s, SE = 7s), $p > .99$. However, both high school students and younger adults were significantly faster at correctly entering their passwords using an iPod Touch than older adults (M = 75s, SE = 8s), ps < .001. For correctly entered passwords using an iPad Mini, there were no significant differences in mean RTs between high school students (M = 11s, SE = 7s), younger adults (M = 14s, SE = 6s), and older adults (M = 23s, SE = 8s), ps > .73. No other RT effects were significant (ps > .11).

2.2.3 Questionnaire

Although detailed analyses of the responses on the questionnaire are not the focus of the current paper, responses to two questions merit discussion. There were significant differences among groups on which type of password they thought was more fun to use, $\chi^2(2) = 11.30$, $p = .004$. Specifically, 26 out of 42 high school students[5] (i.e., 62%), 40 out of 48 younger adults (i.e., 83%), and 18 out of 37 older adults[6] (i.e., 49%) thought that game-based passwords were more fun to use than traditional passwords. There were no differences among groups on which type of password they thought would be easier to remember, $\chi^2(2) = 2.41$, $p = .30$. Interestingly, 88 out of 125 participants[7] (i.e., 71%) thought that traditional passwords would be easier to remember than game-based passwords.

2.3 Discussion

An overall mean PC of 77% is reasonable, given that (a) this is a brand new system, (b) we looked at three different age groups, (c) participants had created a new password, and (d) participants were not specifically told when they were creating their passwords that they would be asked to remember (and re-enter) their passwords. The current results provide evidence that high school students, younger adults, and older adults can all remember game-based passwords

[5] One high school student did not make a clear choice as to which password type he or she thought would be more fun to use.

[6] Three older adults did not make a clear choice as which password type they thought would be more fun to use.

[7] Six participants did not make a clear choice as to which password type they thought would be easier to remember.

fairly well. The fact that older adults and high school students do fairly well is important for password security - and technology more broadly - because research has not paid enough attention to older adults or high school students and our accuracy results are encouraging. Additionally, although high school students and younger adults were significantly faster at correctly entering their passwords using an iPod Touch compared to older adults, this was not the case when using an iPad Mini.

Although, as mentioned previously, cognitive psychologists consider anything beyond approximately 20 seconds long-term memory, it is important to understand whether users can remember their passwords over a longer period of time. Furthermore, it is important to know how participants will perform when asked to remember more than just one password. Finally, since Tao and Adams (2008) report a ~13% improvement in accuracy rate in performance over a 13-week period (78%) compared to performance during the first week (~65%), we examined whether such a practice effect would occur each time a new game is used or whether some of the practice would generalize. That is, perhaps some of this practice effect may be due to learning to use a game-based password in general, such that performance may not return to baseline even after having to switch to a new game. The purpose of Experiment 2 was to examine each of these important issues – memory for game-based passwords over a longer period of time (10 weeks), memory for multiple (five) game-based passwords, and memory after participants are asked to create new passwords and in a new game (that is, switching from chess to Monopoly or vice versa).

3. Experiment 2

3.1 Method

3.1.1 Participants

Thirty-eight participants (30 females) were recruited from the Cleveland State University community between 18 and 50 years old (*M* = 23 years, *SE* = 7 years). However, only 33 participants (25 females) with a mean age of 24 years (*SE* = 8 years) completed all 10 weeks of the experiment. Participants received $100 for their participation, once they completed all 10 weeks of the experiment per our approved Institutional Review Board protocol.

3.1.2 Materials and Procedure

Participants were asked to return to the lab two days per week for a total of 10 weeks.[8] During the first session, participants were first asked to complete a consent form. Participants were then given either an iPod Touch or an iPad Mini with either a chessboard or a Monopoly board application open. Examples of the chessboard and Monopoly boards are shown in Appendices B and C. Participants were asked to create five different passwords either using two or four game pieces. Although participants created – and later entered – all passwords on the same device (either an iPod Touch or an iPad Mini), the different passwords corresponded to five different mock accounts - bank, cell phone, email, laptop, and work computer. After participants created all five passwords, there was a 15-20 minute filled delay during which participants viewed a TED talk. After the delay, participants were instructed to re-enter their passwords on the same device used to create their passwords. Passwords for each of the five different mock

[8] There was one exception. During week nine, one participant entered his or her passwords once, but during week 10, he or she entered his or her passwords three times.

accounts were requested (e.g., “Now enter your bank password”). Participants had up to three chances to correctly enter each password. Participants were then asked to re-enter their passwords for a second time.

During the 10 weeks, participants were asked to enter each of their five different game-based passwords a total of 24 times (for the first two weeks, participants re-entered their five passwords a second time). Each time participants were asked to enter each of their five passwords, the order of presentation was randomized (in order to prevent participants from only learning the order that they entered their five passwords). After entering their five passwords a total of 20 times, participants were asked to create new passwords and they changed games (e.g., from chess to Monopoly or from Monopoly to chess). After participants created five new passwords, there was a 15-20 minute delay during which they were asked to complete the same questionnaire from Experiment 1 about games and passwords and viewed a different TED talk from the first session. After the delay, participants were instructed to re-enter their newly created passwords. Participants had up to three chances to correctly enter each of their passwords. RTs (to correct entries) and PCs were recorded for each participant. RTs were measured from the onset of participants’ initial game piece movement until the *Check Password* button was clicked.

3.1.3 Design

Password Size (two or four pieces), Game (Chess then Monopoly or Monopoly then Chess), and Device (iPod Touch or iPad Mini) were all completely counterbalanced across participants. Thus, we used a 2 × 2 × 2 completely between participants design.

3.2 Results

Two separate 2 (Password Size: two, four) × 2 (Game: Chess than Monopoly, Monopoly than Chess) × 2 (Device: iPod Touch, iPad Mini) × 24 (Entries) mixed ANOVAs were performed, one on mean PCs within three attempts and one on mean RTs to correct attempts.[9]

3.2.1 Accuracy

The overall mean PC was 82% (*SE* = 4%). The main effect of Entry was significant, $F(4.80, 119.92) = 8.73$, $MSE = 1423.37$, $p < .001$, $\eta_p^2 = .26$. The mean PC as a function of entry within three attempts is displayed in Figure 1. The first entry ($M = 54\%$) was significantly less accurate than the 21st entry (i.e., after creating new passwords and switching the game; $M = 72\%$), $p = .024$, demonstrating that there is indeed a carryover effect such that practice using game-based passwords extends to performance to new games.

The main effect of Game was not significant, $F(1, 25) < 1$, $MSE = 11598.38$, $p = .51$, $\eta_p^2 = .02$. The mean PCs for Chess and Monopoly were 85% and 80%, respectively.

The main effect of Password Size was not significant, $F(1, 25) < 1$, $MSE = 11598.38$, $p = .74$, $\eta_p^2 < .01$. The mean PCs for two and four pieces were 84% and 81%, respectively, demonstrating that practice with multiple passwords over an extended period of time results in performance with four piece passwords being equivalent to performance with two piece passwords.

[9] For entries five and six, an iPad Mini participant used an iPod Touch. Also, for the 21st entry, an iPod Touch participant used an iPad Mini.

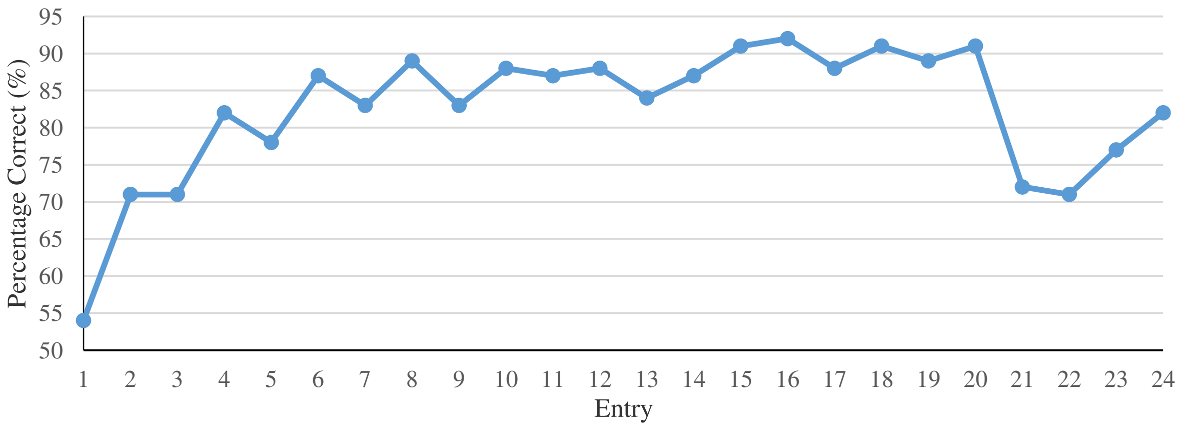


Figure 1. Mean percentage correct (PC) as a function of entry within three attempts.

The main effect of Device was not significant, *F*(1, 25) < 1, *MSE* = 11598.38, *p* = .66, η_p^2 = .01. The mean PCs for iPod Touch and iPad Mini were 8f1% and 84%, respectively. No other effects of PC were significant (all *p*s > .10).

We also explored the effect of Account Type (bank, cell phone, email, laptop, and work computer) on accuracy in Experiment 2. There was no main effect of Account Type, nor did Account Type interact with Password Size, Game, or Device, all *p*s > .23. The mean PCs as a function of the five password types appear in Table 3.

Table 3

Mean Percentages Correct (PCs), Reaction Times (RTs), and Standard Errors (SEs) as a Function of Account Type

	PC		***RT***	
Account	**M**	**SE**	**M**	**SE**
Bank	85	4	10	1
Cell Phone	79	6	10	1
Email	81	4	11	1
Laptop	82	4	10	1
Work Computer	85	3	11	1

3.2.2 Reaction Time

The overall mean RT to correctly enter a password was 11s (*SE* = 1s). The main effect of Entry is significant, $F(5.49, 137.25) = 6.41$, $MSE = 71.55$, $p < .001$, $\eta_p^2 = .20$. The mean RT as a function of entry within three attempts is displayed in Figure 2. The first entry (*M* = 17s, *SE* = 2s) took significantly longer than the 21st entry (i.e., after creating new passwords and switching the game; *M* = 9s, *SE* = 1s), $p < .001$.

The main effect of Game was not significant, $F(1, 25) < 1$, $MSE = 265.99$, $p = .62$, $\eta_p^2 = .01$. The mean RTs for Chess and Monopoly were both 11s. The main effect of Password Size was significant, $F(1, 25) = 19.75$, $MSE = 265.99$, $p < .001$, $\eta_p^2 = .44$. Participants with passwords consisting of two pieces (*M* = 8s, *SE* = 1s) were significantly faster at correctly entering their password than participants with passwords consisting of four pieces (*M* =13s, *SE* = 1s).

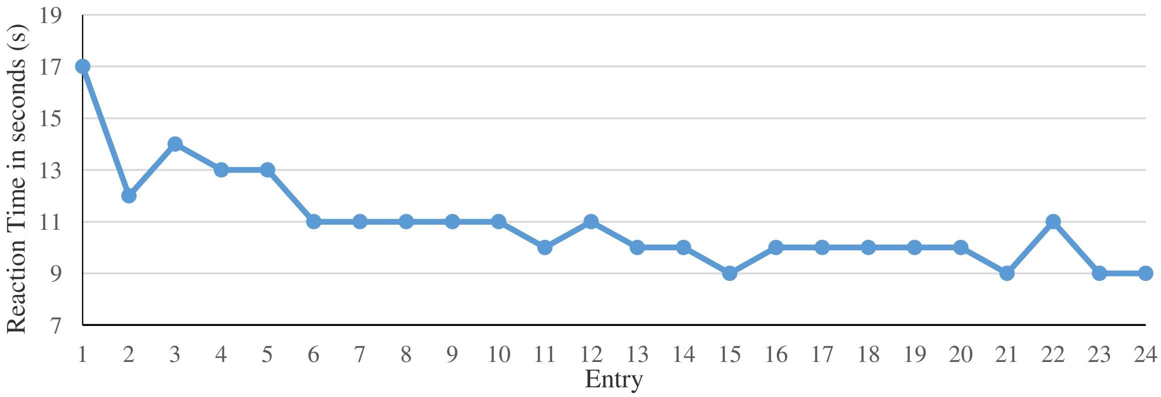


Figure 2. Mean reaction time (RT) in seconds (s) as a function of entry within three attempts.

The main effect of Device was significant, $F(1, 25) = 41.05$, $MSE = 265.99$, $p < .001$, $\eta_p^2 = .62$. Participants using an iPod Touch (*M* = 15s, *SE* = 1s) to enter their passwords were significantly slower at correctly entering their passwords than participants using an iPad Mini (*M* = 7s, *SE* = 1s).

The Entry × Device interaction was marginally significant, *F*(4.80, 119.92) = 2.20, *MSE* = 1423.37, *p* = .05, η_p^2 = .08. The mean RT as a function of Entry and Device within three attempts is displayed in Figure 3. Simple effects for Entry were preformed. For the first entry, the iPod Touch (*M* = 24s, *SE* = 2s) was significantly slower than the iPad Mini (*M* = 11s, *SE* = 2s), *p* < .001. For the 21st entry, the iPod Touch (*M* = 11s, *SE* = 1s) trended to be slower than the iPad Mini (*M* = 7s, *SE* = 1s), *p* =.088. No other RT effects were significant (*ps* > .12).

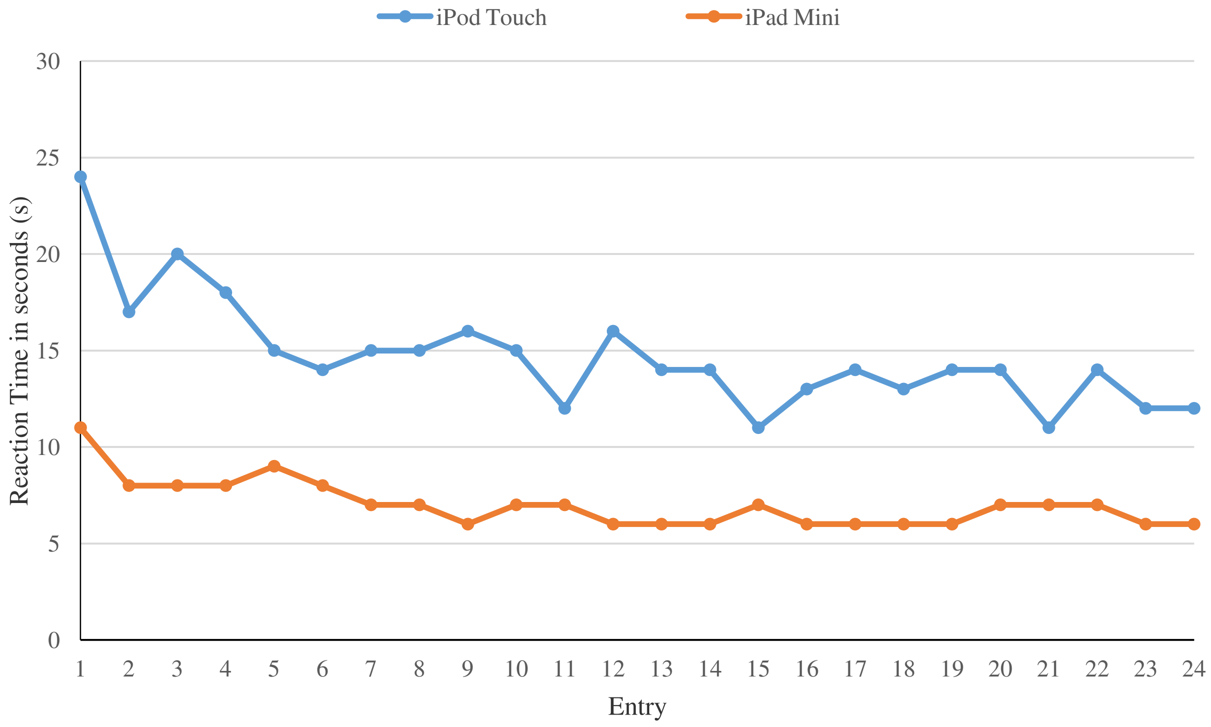


Figure 3. Mean reaction time (RT) in seconds (s) as a function of entry and device within three attempts.

We also explored the effect of Account Type (bank, cell phone, email, laptop, and work computer) on RT in Experiment 2. There was no main effect of Account Type, nor did Account

Type interact with Password Size, Game, or Device, all *p*s > .18. The mean RTs as a function of the five password types appear in Table 3.

3.2.3 Questionnaire

30 out of 32 participants[10] (i.e., 91%) thought that game-based passwords were more fun to use than traditional passwords. However, 23 out of 31 participants[11] (i.e., 74%) thought that traditional passwords would be easier to remember than game-based passwords.

3.3 Discussion

An overall mean PC of 82% for participants remembering multiple game-based passwords over 10 weeks is reasonable. Initially, participants' accuracy for multiple game-based passwords was poor (i.e., 54%), but performance quickly improved with practice. Moreover, when participants changed games and created new passwords, their performance was significantly better than the first time they entered their passwords.

4. General Discussion

Based on our empirical data, the Game Changer Password System is promising. Mean PCs of 77% for Experiment 1 and 82% for Experiment 2 are reasonable given that this is a brand new system. Furthermore, these mean PCs include the iPod Touch in which the game boards

[10] One participant did not make a clear choice as to which password type he or she thought would be more fun to use.

[11] Two participants did not make a clear choice as to which password type they thought would be easier to remember.

were not completely visible on the screen, and thus had to be moved around. Although these PCs are comparable to what has been reported with Pass-Go by Tao and Adams (2008), in which login success rate was 78%, presumably performance would be even higher when the application is suited to the device, participants have more familiarity using a game-based system, and they are using a password they actually use to log in to a device or access an account, and thus have a personal motivation to try and remember. Also, 77% and 82% are the means, as indicated in Table 1, performance was as high as 100% in some conditions. Also, our analysis of the types of errors participants made in Experiment 1 leads to an important conclusion. A system that required all correct pieces and all correct locations, but allowed any combination of pieces and locations would have increased accuracy by over 13% (averaged across all three age groups). These results pave the way for future research into the use of game-based passwords – including the possibility of developing more flexible systems (e.g., systems that are more or less "forgiving" depending on the needed underlying security).

Mean RTs of 28s for Experiment 1 and 11s for Experiment 2 are likely longer than most users would put up with to, say, access their smartphone. However, 28s in Experiment 1 is not too bad considering this is averaging across three age groups, in which some participants may not have had experience with using the device to log in. Additionally, 11s in Experiment 2 is promising since this is averaging across multiple passwords over 10 weeks. In addition, these mean RTs of 28s and 11s are nowhere near the up to three minutes that has been reported for another graphical system (Brostoff & Sasse, 2000) and is even satisfactory when compared to the mean login time of 34s reported for some alphanumeric passwords (Vu et al., 2007). Furthermore, many participants looked over their password to check for accuracy before clicking the *Check Password* button that stopped the clock. Future investigations may want to stop the clock when the last piece is placed on the board (presumably when an end-users' entry of the password could be complete). Also, 28s and 11s are the means, as indicated in Figure 3, performance was as fast as 5s in some conditions.

We should mention that the version of chess used by Pandey (2012) required users to remember two rules, one involving the bishop and the other involving the rook. Such a system may be less appealing to end-users, particularly users that find it difficult to learn or remember such rules, such as young children and older adults, respectively. Also, the Snakes and Ladders game-based system reported by Malempati and Mogalla (2011) was based on users remembering their favorite moves. Thus, unlike the Game Changer Password System, their system required users to know how to play the game.

The types of passwords chosen – in particular the number of times each piece and each location – were chosen is summarized in the figures of the games that appear in Appendices B and C[12]. A breakdown of the passwords chosen for Experiment 2 only as a function of the five different mock interfaces (bank, cell phone, email, laptop, and work computer) appear in Appendices D and E. It does not appear that the pieces or the locations are equally probable. Consequently, security calculations may need to be revisited to determine effective password space. However, given the relatively small sample size and the number of passwords created, such a conclusion may be premature. Nevertheless, even if the pieces or locations are not equally probable, these data are encouraging in that they demonstrate a good degree of variability – both in the pieces and the locations chosen, and in both chess and Monopoly. All of the pieces in both games were selected as part of the password at least once, and many different

[12] These photos represent the largest sample we have available, including all data collected for Experiments 1 and 2 as well as other ongoing projects.

locations were chosen as well. Nevertheless, some pieces and locations stand out as having been chosen with relatively higher probability.

4.1 Limitations

4.1.1 Sample

The data reported in this paper may not be completely representative of the general population. Moving forward, the Game Changer Password System should be tested with significantly larger and more diverse samples.

4.1.2 Data Analyses

The current report is based on data recently collected. Additional analyses are planned, for example, in order to examine the relationships between piece and location choice (e.g., Is the thimble in Monopoly more likely to be put in jail than other pieces? Is a black king in chess more likely to be put in a particular location than other pieces?), between piece and piece (e.g., If a user selects a thimble is he or she more likely to select a dog as one of his or her other pieces?), and between location and location (e.g., If a user selects a railroad as one of his or her locations, is he or she more likely to select another railroad as a location of another of his or her pieces?), as well as the relationships between reported experience, say with chess, and password choice and performance.

Although we have a rich data set that will continue to be explored for these and other such possible relationships, the current data make important contributions. That is, based on the analyses we have reported here, we now know 1) that participants can remember passwords fairly well – across different games, password sizes, and on different devices, 2) that participants choose a variety of pieces and locations to create their passwords, 3) that, not surprisingly, most of the participants report game-based passwords as being more fun to use than alphanumeric passwords, and 4) that participants report traditional alphanumeric passwords as being easier to remember than game-based passwords. This last finding is not surprising since other researchers have found a lack of correlations between "perceived" and "real" risks of shoulder-surfing between certain graphical and alphanumeric password entries (Tari et al., 2006), and other researchers report that participants' perceptions of password security do not always match reality (Ur, et al., 2016). Despite the incredible leap in security over traditional alphanumeric-based systems, these data suggest that any implementation of a game-based system would require some education regarding the underlying calculations and associated higher degrees of security.

4.1.3 Ecological validity

The passwords we collected did not protect real user accounts, a common limitation in password research. The game boards did not completely display on the iPod Touch, requiring the participants to move the board around as necessary. Although this was done intentionally to determine whether or not this unpleasant (not user-friendly) aspect of the system would affect memory performance, obviously a real implementation for end-users would be done in a way that adjusts for screen size, device, etc. Although testing took place using two different devices, these devices did not belong to the participants, and the participants may or may not own, or have experience using, the device on which they were tested. Moving forward, the Game Changer Password System should also be tested on more devices and to access real accounts. For example, we report in Experiment 2 that there was no main effect of Account Type on accuracy, nor did Account Type interact with Password Size, Game, or Device. However, these null effects may be due the "hypothetical systems" in our experimental tests. Such differences may emerge when evaluating performance in which these game-based passwords are used to access real accounts.

4.1.4 System Implementation

Currently, most authentication systems require a combination of a username and a password with alphanumeric characters. Any graphical password scheme, including the Game Changer Password System, would require a redesign of user interface for existing systems. Although such redesigns are likely not trivial, requiring substantial investments, this limitation, of course, only applies to existing systems, and not to new systems designed for a game-based approach. Also, any such redesigns may be worthwhile if they enhance the underlying security.

4.2 Future Work

4.2.1 Additional testing

Additional testing should include more extensive experimental and non-experimental evaluations of the Game Changer Password System using a fully functional and realistic Netflix-style interface. One of the strengths of the current study is that data were collected from older adults and high school students. As mentioned earlier, password security research - and technology studies more broadly - has not paid sufficient attention to these populations of users. Future work might consider extending to even younger children, given that children are using devices at an increasingly young age (Rideout, Foehr, & Roberts, 2010). In fact, younger children (and others) might benefit from learning about game-based passwords by using password games, as has been done with teaching about the importance of passwords more generally (Gardner & Atkinson, 2012).

4.2.2 Traditional Passwords

The data reported here are not compared to traditional passwords. Future work will examine how these game-based passwords compare to traditional passwords. Although doing so may seem to be sufficiently easy, this is not necessarily the case. If we wanted to allow users to choose their own traditional and game-based passwords, we could not prevent users from re-using a traditional password that they currently already use (in which case obtained data would not lead to a fair comparison). On the other hand, if users are assigned their traditional and game-based passwords, then we take away the personalization the user has, and it is unclear what the consequence of doing so might have for participants' memory for either type of password. Nevertheless, a previous study compared text passwords and PassPoints, a click-based graphical password system. The researchers found that in a recall condition (Recall-2) in which participants were tested 12 to 15 days after initially creating their passwords, participants' ("primarily university students") mean recall performance within three attempts for six text passwords (8-character minimum) was 59%, and participants' mean recall performance within three attempts for six PassPoint passwords (click on five different click-points) was 57% (Chiasson, Forget, Stobert, van Oorschot, & Biddle, 2009). There are a number of differences between this previous study and our current experiments, and consequently any comparisons should be made with these differences in mind. Despite this important caveat, it is encouraging to note that participants' overall mean performance was 82% in Experiment 2 for the multiple (five) game-based passwords over 10 weeks.

Recent work has demonstrated that textual passwords are particularly vulnerable to targeted online guessing (Wang, Zhang, Wang, Yan, & Huang, 2016). Textual passwords may be influenced by personal information and textual passwords may be reused, both of which contribute to the vulnerability of textual passwords to online guessing. Although it seems unlikely that the Game Changer Password System is vulnerable to such targeted online guessing, future tests could be designed to answer this question.

4.2.3 Implementations

Although our current version of the Game Changer Password System is perfectly fine for the scientific investigations reported in the current paper, we plan to develop a system that is more realistic and user friendly, especially across multiple devices. Doing so would not only allow for new longer-term tests with increased ecological validity, but it would also allow us to adequately critique a more detailed instantiation of the Game Changer Password System by the Usability-Deployability-Security (UDS) evaluation framework as it was originally intended by Bonneau et al. (2012). The UDS evaluation framework was developed by Bonneau et al. (2012) in order to compare the strengths and weaknesses of all user authentication schemes. Perhaps not surprisingly, no existing password system can meet all of the stringent 25 tests that are known concerns for password security. Thus, it has proven impossible for any existing end-user authentication system to satisfy all of these demands. Bonneau et al. (2012) therefore comment: "Replacing passwords with any of the schemes examined is not a question of giving up an inferior technology for something unarguably better, but of giving up one set of compromises or tradeoffs in exchange for another" (p. 562). When doing so, we will consider recent work exploring the design space of graphical passwords on electronic devices (Schaub, Walch, Könings, & Weber, 2013).

Following the work of Schechter, Herley, and Mitzenmacher (2010), a system for developing password policies could be devised in which certain passwords are blacklisted – for example, once a threshold of confusability, based on empirical data, is crossed, or for password pieces, locations, or combinations (of pieces and locations) that are used too frequently. A recently proposed algorithm may be particularly efficient because it is adaptive by continually tracking passwords and, when a password becomes too popular, the password could be rejected for new users - existing users can receive a warning to change their password (Chanda, 2016). A similar system could be developed for game-based passwords. We plan to continue to examine users' choice of games, game pieces, locations, and various relationships (e.g., between game pieces and locations). Doing so will allow us to gain a more complete understanding of the role that user choice plays in various game-based passwords, as has been done with other graphical passwords (Nali & Thorpe, 2004), and more recently with real passwords (Shen, Yu, Xu, Yang, & Guan, 2016).

5. Conclusion

The Game Changer Password System is a new way of thinking about password security (Taneski, 2014), one that is informed by cognitive psychology. Storing passwords in game positions offers an alternative to the bane of alphanumeric systems that are hard for users to buy into because they require awkward memorization. The practical need for users to remember multiple passwords and to change existing passwords periodically (although see Zhang-Kennedy, Chiasson, & van Oorschot, 2016) leads to a blurring of passwords, user frustration, and system breakdown as users defect.

As stated earlier, we also believe that the Game Changer Password System emphasizes two new factors: fun and relative usability in different populations. We hypothesize that the game approach will appeal to anyone who likes games in general and that the approach might be an olive branch to populations intimidated – or just put off – by the stringent demands of password schemes. We hypothesize specifically that older populations will prefer the Game Changer Password System to alphanumeric codes and that all populations will have less errors following a change in password, as long as they store the new password in a different game. Recall that the Picture Superiority Effect refers to the finding that memory is typically better for pictures than for other types of information, including words. There is evidence that older adults would especially benefit from the Picture Superiority Effect (Park, Puglisi, & Sovacool, 1983). It is certainly interesting that older adults' performance was significantly better – in terms of PC – than high school students' performance. Finding that older adults outperform a group of high school students in a task with a strong memory component may be atypical, and suggests real potential in having such a game-based password system adopted by older adults.

We envision game changing as an element in cybersecurity that operates most efficiently when users *toggle* between multiple passwords, each of which needs to be mathematically robust. Insight from cognitive psychology shows that switching games minimizes the likelihood of memory interference. The use of graphical passwords is gaining momentum (Renaud, Mayer, Volkamer, & Maguire, 2013; Jenkins, McLachlan, & Renaud, 2014). Our approach builds on this momentum and capitalizes on the many strengths of a graphical approach, while avoiding many of the reported challenges (Gao, Jia, Ye, & Ma, 2013; Gao, Ma, Jia, & Ye, 2012; Zakaria, Griffiths, Brostoff, & Yan, 2011).

Although we have argued above that proper objective evaluation in the UDS framework requires a more detailed instantiation of our system, we nevertheless point out that nothing about our approach prevents flexibility in how the system is implemented. Indeed, we argue that such flexibility is one of the system's major strengths. Stacking, layering, the number of pieces, and other modifications increase the security of the password. The desired level of security will likely depend on whether you are using the system to access your smartphone or your online bank account, to withdraw money at an ATM machine (Moncur & Leplâtre, 2007), to enter the password to your home alarm system, enter (or start) your vehicle, enter your office (or the office building), or, as a manager at a local bank, to open the bank vault. As you can see, the Game Changer Password System not only has the potential to replace (or complement) existing password schemes for electronic devices, but it also has the potential to replace (or complement) the use of keys.

Acknowledgements

This paper is based upon research supported by the United States National Science Foundation under Grant No: 1343141. Portions of this work were presented at the 27th Association for Psychological Science Annual Convention in New York, N.Y. and at the 28th Association for Psychological Science Annual Convention in Chicago, IL.

References

Alkaldi, N. & Renaud, K. (2016). Why do people adopt, or reject, smartphone password managers? *Proceedings of the European Workshop on Usable Security (EuroUSEC).*

Biddle, R., Chiasson, S., & van Oorschot, P. C. (2012). Graphical passwords: Learning from the first twelve years. *ACM Computing Surveys (CSUR)*, *44*(4), 19.

Bonneau, J., Herley, C., van Oorschot, P. C., & Stajano, F. (2015). Passwords and the evolution of imperfect authentication. *Communications of the ACM, 58,* 78-87.

Bonneau, J., Herley, C., van Oorschot, P. C., & Stajano, F. (2012). The quest to replace passwords: A framework for comparative evaluation of web authentication schemes. *Security and Privacy (SP), 2012 IEEE Symposium on (pp. 553-567). IEEE.*

Brostoff, S., & Sasse, M. A. (2000). Are Passfaces more usable than passwords? A field trial investigation. *People and Computers XIV—Usability or Else!* (pp. 405-424). Springer London.

Brumen, B. & Cernezel, A. (2014). Brute force analysis of PsychoPass-generated Passwords. *Information and Communication Technology, Electronics, and Microelectronics (MIPRO), 37th International Convention on Opatija,* 1366-1371.

Brumen, B. & Taneski, V. (2015). Moore's curse on textual passwords. *Information and Communication Technology, Electronics, and Microelectronics (MIPRO), 38th International Convention on Opatija,* 1360-1365.

Chanda (2016). Password security: An analysis of password strengths and vulnerabilities. *International Journal of Computer Network and Information Security, 7,* 23-30.

Cheswick, W. (2013). Rethinking passwords. *Communications of the ACM*, *56*(2), 40-44.

Chiasson, S., Forget, A., Stobert, E., van Oorschot, P.C., & Biddle, R. (2009). Multiple password interference in text passwords and click-based graphical passwords. Proceedings of CCS, ACM, New York, 500-511.

Curran, K. & Snodgrass, A. (2015). A novel cue based picture word shape character password creation scheme. *International Journal of Digital Crime and Forensics, 7,* 37-59.

De Angeli, A., Coventry, L., Johnson, G., & Renaud, K. (2005). Is a picture really worth a thousand words? Exploring the feasibility of graphical authentication systems. *International Journal of Human-Computer Studies*, *63*(1), 128-152.

Eich, J., John, L., Smith, K., & Cankaya, E.C. (2016). Extended password security via cloud: CloudPass. In D. Nicholson (Ed.), *Advances in Human Factors in Cybersecurity, Advances in Intelligent Systems and Computing,* 33-42.

Engelmann, M. (2009). Competition between two memory traces for long-term recognition memory. *Neurobiology of Learning and Memory, 91,* 58-65.

Gao, H., Jia, W., Ye, F., & Ma, L. (2013). A survey on the use of graphical passwords in security. *Journal of Software, 8*, 1678-1698.

Gao, H., Ma, L., Jia, W., & Ye, F. (2012). Multiple Password interference in graphical passwords. *International Journal Information and Computer Security, 5*, 11-27.

Gardner, R., & Atkinson, S. (2012). E–learning and Password Games. *Advances in Communications, Computing, Networks and Security Volume 9*, *8*, 95.

Grosse, E., & Upadhyay, M. (2013). Authentication at scale. *Security & Privacy, IEEE*, *11*(1), 15-22.

Gutmann, A., Volkamer, M., & Renaud, K. (2016). Memorable and secure: How do you choose your PIN? *Proceedings of the Tenth International Symposium on Human Aspects of Information Security & Assurance.*

Herley, C., & van Oorschot, P. (2012). A research agenda acknowledging the persistence of passwords. *Security & Privacy, IEEE*, *10*(1), 28-36.

Jenkins, (2014). Facelock: Familiarity-based graphical authentication. *PeerJ 2:e444; DOI 10.7717/peerj.444.*

Ji, S., Yang, S., Hu, X., Han, W., Li, Z., & Beyah, R. (2016). Zero-sum password cracking game: A large-scale empirical study on the crackability, correlation, and security of passwords. *IEEE Transactions on Dependable and Secure Computing, 99,* 1545-5971.

Khan, W. Z., Aalsalem, M. Y., & Xiang, Y. (2011). A graphical password based system for small mobile devices. *International Journal of Computer Science Issues, 8*, 145-154.

Lo, C.C-W. (2016). Empirical study of secure password creation habit. *Foundations of Augumented Cognition: Neuroergonomics and Operational Neuroscience, 9744*, 189-197.

Malempati, S. & Mogalla, S. (2011). An ancient indian board game as a tool for authentication. *International Journal of Network Security & Its Applications, 3*, 154-163.

Manning, P., M^{c}Lennan, C.T., & Zhu, Y. *Authentication Method for a Computing Device Using Interactive Game Board and Game Piece Images.* Patent U.S. Provisional Application. Application No.: 61782062.

Moncur, W., & Leplâtre, G. (2007). Pictures at the ATM: Exploring the use of multiple graphical passwords. *CHI 2007 Proceedings*, San Jose, CA, 887-894.

Nali, D., & Thorpe, J. (2004). Analyzing user choice in graphical passwords. *School of Computer Science, Carleton University, Tech. Rep. TR-04-01.*

Paivio, A., & Csapo, K. (1973). Picture superiority in free recall: Imagery or dual coding? *Cognitive Psychology, 5,* 176 –206.

Pandey, S. K. (2012). Chess game as a tool for authentication scheme. *International Journal of Scientific Research Engineering & Technology, 1*, 076-083.

Park, D.C., Puglisi, J.T., & Sovacool, M. (1983). Memory for pictures, words, and spatial location in older adults: Evidence for pictorial superiority. *Journal of Gerontology, 38,* 582-588.

Ramapriya, M., Yamini, R., Moorthy, M.K., & Scholar, P.G. (2015). Captcha as graphical passwords: A new security primitive based on hard AI problems. *SSRG International Journal of Mobile Computing & Applications, 2,* 34-38.

Renaud, K., Mayer, P., Volkamer, M., & Maguire, J. (2013). Are graphical authentication mechanisms as strong as passwords? *Proceedings of the 2013 Federated Conference on Computer Science and Information Systems*, 837-844.

Rideout, V. J., Foehr, U. G., Roberts, D. F. (2010). Generation M2: Media in the Lives of 8- to 18-Year -olds. *A Kaiser Family Foundation Study.* Menlo Park, California; Henry J. Kaiser Family Foundation.

Rao, M.K., Pravallika, C.V., Priyanka, G., & Kumar, M. (2016). *Innovations in Computer Science and Engineering, 413,* 105-112.

Schaub, F., Walch, M., Könings, B., & Weber, M. (2013). Exploring the design space of graphical passwords on smartphones. *Symposium on Usable Privacy and Security 2013,* Newcastle, UK.

Schechter, S., Herley, C., and Mitzenmacher, M. (2010). Popularity is everything: A new approach to protecting passwords from statistical-guessing attacks. *Proc. HotSec'10.*

Shay, R., Bauer, L., Christin, N., Cranor, L.F., Forget, A., Komanduri, S., Mazurek, M.L., Melicher, W., Segreti, S.M., & Ur, B. (2015). In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems,* 2903-2912.

Shen, C., Yu, T., Xu, H., Yang, G., & Guan, X. (2016). User practice in password security: An empirical study of real-life passwords in the wild. *Computers & Security, 61,* 130-141.

Suo, X., Zhu, Y., & Owen, G.S. (2005). Graphical passwords: A survey. *Annual Computer Security Applications Conference (ACSAC).*

Taneski, V. Hericko, M., & Brumen, B. (2014). Password security – No change in 35 years? *Information and Communication Technology, Electronics, and Microelectronics (MIPRO), 37th International Convention on Opatija,* 1360-1365.

Tao, H. & Adams, C. (2008). Pass-Go: A proposal to improve the usability of graphical passwords. *International Journal of Network Security, 7*, 273-292.

Tari, F., Ozok, A., & Holden, S. H. (2006, July). A comparison of perceived and real shoulder-surfing risks between alphanumeric and graphical passwords. *Proceedings of the second symposium on Usable privacy and security* (pp. 56-66). ACM.

Underwood, B.J. (1957). Interference and forgetting. *Psychological Review, 64*, 49-60.

Ur, B., Bees, J., Segreti, S.M., Bauer, L., Christin, N., & Cranor, L.F. (2016). Do users' perceptions of password security match reality? In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems,* 3748-3760.

US Department of Health and Human Services Administration on Aging (2010). *Aging Statistics.* Retrieved from http://www.aoa.acl.gov/Aging_Statistics/index.aspx van Oorschot, P.C., Salehi-Albari, A., and Thorpe, J. (2010). Purely automated attacks on Passpoints-style graphical passwords. *Informaiton Forensics and Security, IEEE, 5,* 393-405.

Vitevitch, M.S. & Luce, P.A. (1998). When words compete: Levels of processing in perception of spoken words. *Psychological Science, 9,* 325-329.

Vu, K. P. L., & Hills, M. M. (2013). The influence of password restrictions and mnemonics on the memory for passwords of older adults. *Human Interface and the Management of Information. Information and Interaction Design* (pp. 660-668). Springer Berlin Heidelberg.

Vu, K. P. L., Proctor, R. W., Bhargav-Spantzel, A., Tai, B. L. B., Cook, J., & Schultz, E. E. (2007). Improving password security and memorability to protect personal and organizational information. *International Journal of Human-Computer Studies*, *65*(8), 744-757.

Wang, D., He, D., Cheng, H., & Wang, P. (2016). fuzzyPSM: A new password strength meter using fuzzy probabilistic context-free grammars. *46th Annual IEEE/IFIP International Conference on Dependable Systems and Networks.*

Wang, D. & Wang, P. (2016). The Emperor's new password creation policies: An evaluation of leading web services and the effect of role in resisting against online guessing. *Proceedings of the 20th European Symposium on Research in Computer Security (ESORICS).*

Wang, D., Zhang, Z., Wang, P., Yan, J., & Huang, X. (2016). Targeted online password guessing: An underestimated threat. *Proceedings of the 23rd ACM Conference on Computer and Communications Security*, 1-13.

Whalen, T. (2011). Security as if People Mattered. *Security & Privacy, IEEE*, *9*(4), 64-67.

Wiedenbeck, F., Waters, J., Birget, J.-C., Brodskiy, A., Memon, N. (2005). PassPoints: design and longitudinal evaluation of a graphical password system. *International Journal of Human–Computer Studies, 63,* 102–127.

Zakaria, N. H., Griffiths, D., Brostoff, S. & Yan, J. (2011). Shoulder surfing defence for recall-based graphical passwords. *Symposium on Usable Privacy and Security 2011*, Pittsburgh, PA.

Zhang-Kennedy, L., Chiasson, S., van Oorschot, P. (2016). Revisiting password rules: Facilitating human management of passwords. *APWG Symposium on Electronic Crime Research, IEEE,* 81-90.

Zhu, Y., Manning, P., & M^{c}Lennan, C.T. *The Game Changer: A New Model for Password Security.* Grant proposal submitted to, and funded by, the National Science Foundation. Grant No: 1343141.

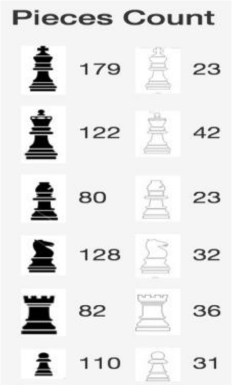
Appendix A

All participants were asked the following questions[13]:

- Do you know how to play chess? (Yes or No)
- How would you rate your experience playing chess? (Low = 1, High = 7)
- Do you know how to play Monopoly? (Yes or No)
- How would you rate your experience playing Monopoly? (Low = 1, High = 7)
- Which do you think would be more fun to use: traditional passwords that use letters, numbers, and/or characters OR game-based passwords like this?
- Which do you think would be easier to remember: traditional passwords that use letters, numbers, and/or characters OR game-based passwords like this?
- Do you have any comments about using game-based passwords like this for smart phones, computers, tablets, ATM machines, online accounts (such as your bank account, etc.?
- Do you have any suggestions on where this password system would be the most useful (e.g., cell phone, laptop, online accounts, etc.)?
- How could this system be tweaked to increase the likelihood of your using this login?

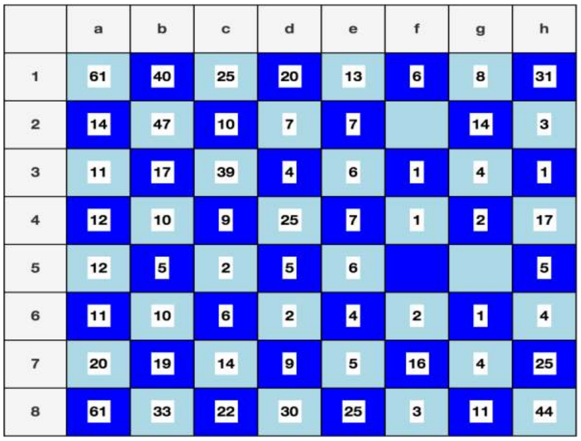
[13] Thirty-two young adults in Experiment 1 did not receive the last two questions.

Appendix B



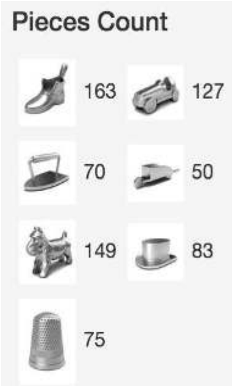
Pieces Count

♚	179	♔	23
♛	122	♕	42
♝	80	♗	23
♞	128	♘	32
♜	82	♖	36
♟	110	♙	31



	a	b	c	d	e	f	g	h
1	61	40	25	20	13	6	8	31
2	14	47	10	7	7		14	3
3	11	17	39	4	6	1	4	1
4	12	10	9	25	7	1	2	17
5	12	5	2	5	6			5
6	11	10	6	2	4	2	1	4
7	20	19	14	9	5	16	4	25
8	61	33	22	30	25	3	11	44

Appendix C



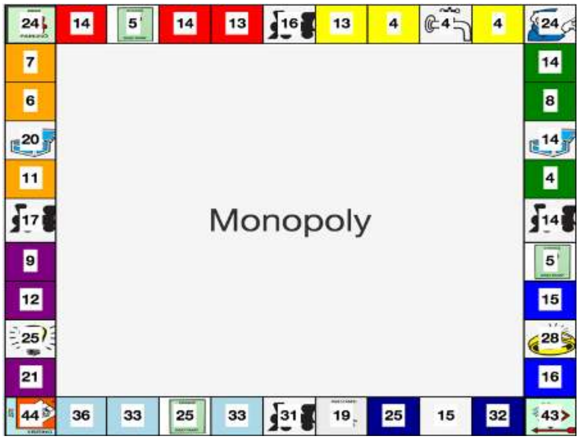
Pieces Count

163 127

70 50

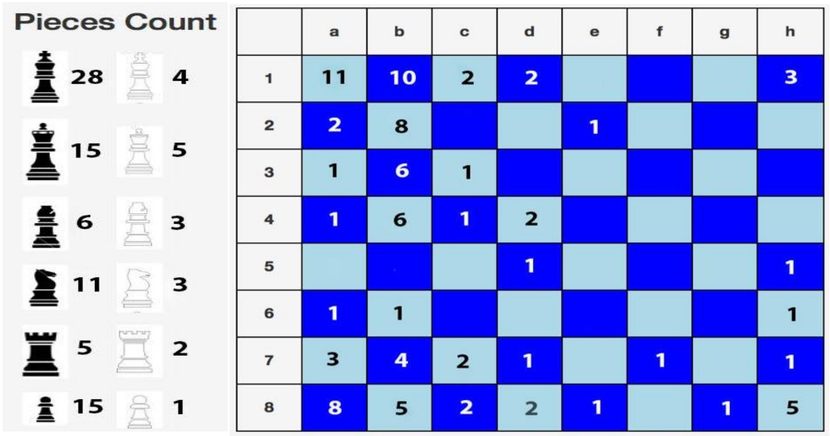
149 83

75



Appendix D

bank

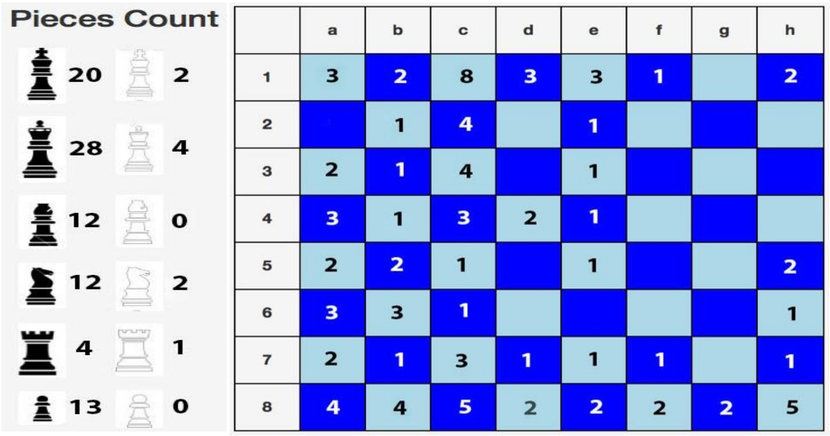


Pieces Count

♚	28	♔	4
♛	15	♕	5
♝	6	♗	3
♞	11	♘	3
♜	5	♖	2
♟	15	♙	1

	a	b	c	d	e	f	g	h
1	11	10	2	2				3
2	2	8			1			
3	1	6	1					
4	1	6	1	2				
5				1				1
6	1	1						1
7	3	4	2	1		1		1
8	8	5	2	2	1		1	5

cell phone

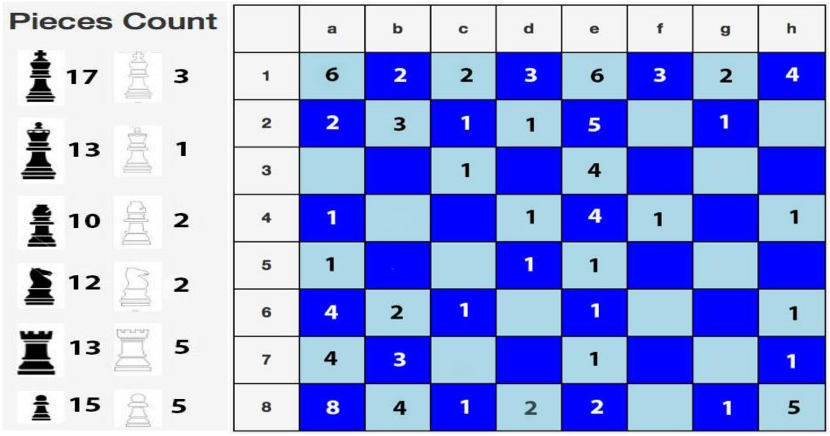


Pieces Count

♚	20	♔	2
♛	28	♕	4
♝	12	♗	0
♞	12	♘	2
♜	4	♖	1
♟	13	♙	0

	a	b	c	d	e	f	g	h
1	3	2	8	3	3	1		2
2		1	4		1			
3	2	1	4		1			
4	3	1	3	2	1			
5	2	2	1		1			2
6	3	3	1					1
7	2	1	3	1	1	1		1
8	4	4	5	2	2	2	2	5

email



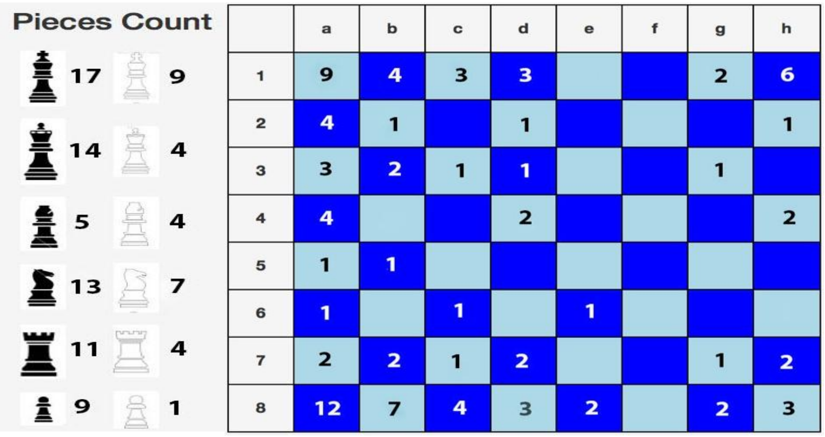
Pieces Count

♚	17	♔	3
♛	13	♕	1
♝	10	♗	2
♞	12	♘	2
♜	13	♖	5
♟	15	♙	5

	a	b	c	d	e	f	g	h
1	6	2	2	3	6	3	2	4
2	2	3	1	1	5		1	
3			1		4			
4	1			1	4	1		1
5	1			1	1			
6	4	2	1		1			1
7	4	3			1			1
8	8	4	1	2	2		1	5

Appendix D continued

laptop

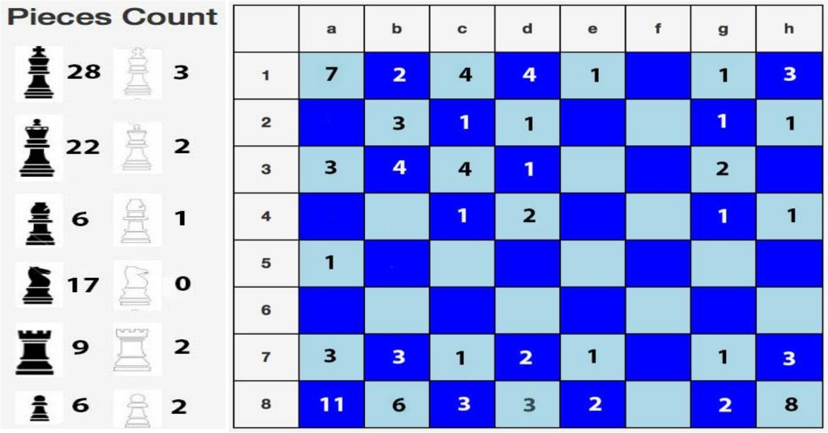


Pieces Count

♚	17	♔	9
♛	14	♕	4
♝	5	♗	4
♞	13	♘	7
♜	11	♖	4
♟	9	♙	1

	a	b	c	d	e	f	g	h
1	9	4	3	3			2	6
2	4	1		1				1
3	3	2	1	1			1	
4	4			2				2
5	1	1						
6	1		1		1			
7	2	2	1	2			1	2
8	12	7	4	3	2		2	3

work computer



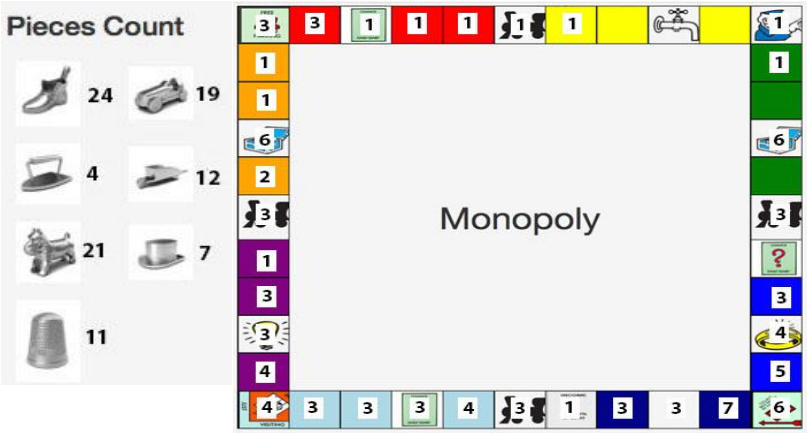
Pieces Count

♚	28	♔	3
♛	22	♕	2
♝	6	♗	1
♞	17	♘	0
♜	9	♖	2
♟	6	♙	2

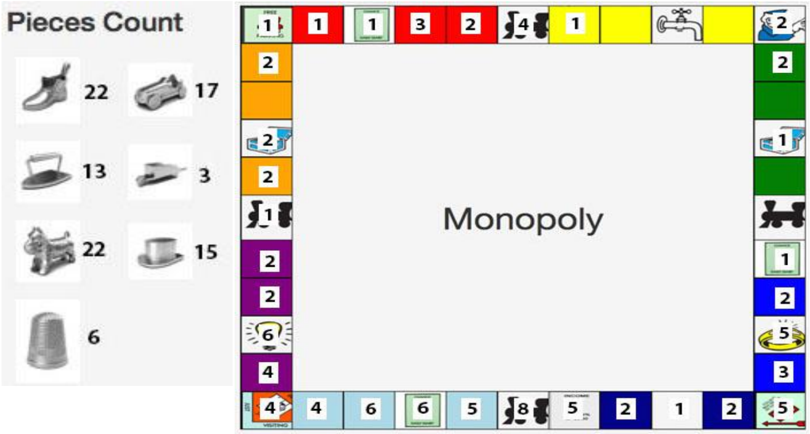
	a	b	c	d	e	f	g	h
1	7	2	4	4	1		1	3
2		3	1	1			1	1
3	3	4	4	1			2	
4			1	2			1	1
5	1							
6								
7	3	3	1	2	1		1	3
8	11	6	3	3	2		2	8

Appendix E

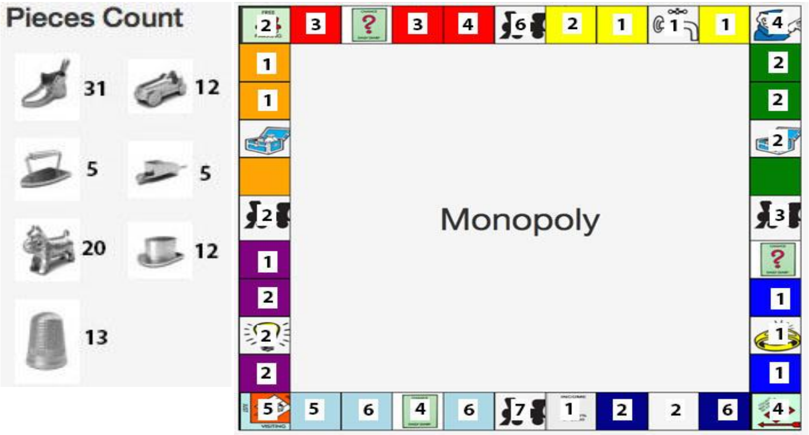
bank



cell phone

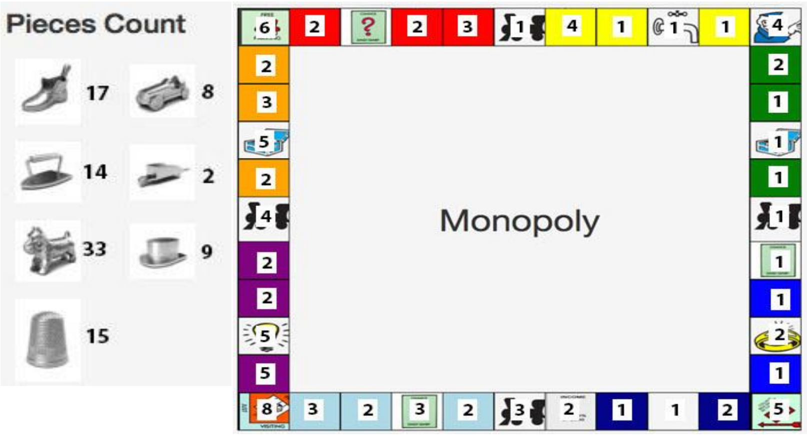


email

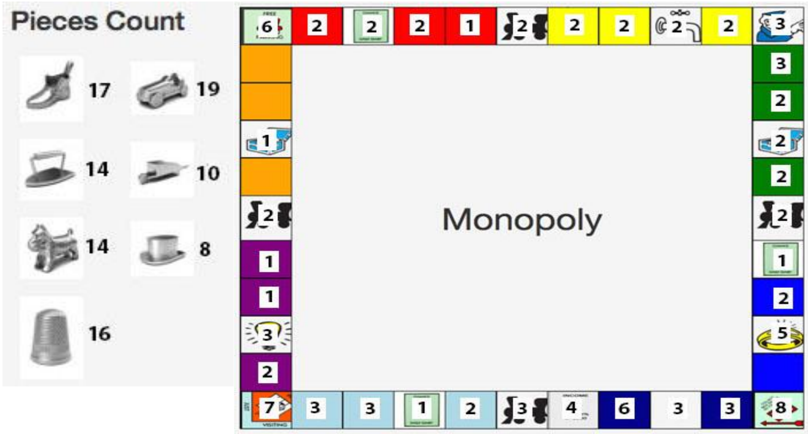


Appendix E continued

laptop



work computer


Pieces Count
17
19
14
10
14
8
16
Monopoly