Personal and Ubiquitous Computing
https://doi.org/10.1007/s00779-021-01571-y

ORIGINAL PAPER

Understanding users' perceptions to improve fallback authentication

Nicholas Micallef[1] · Nalin Asanka Gamagedara Arachchilage[2]

Received: 31 October 2020 / Accepted: 20 April 2021

Abstract
Despite receiving a lot of scrutiny and criticism, security questions are still widely adopted. Although new techniques are continuously being proposed to improve fallback authentication (i.e. security questions design), little research investigated users' security and memorability perceptions. Previous research found that users' perceptions are important because they can impact the adoption of security techniques. Hence, this research contributes to security questions research by investigating (with a study of $n = 30$) how users select security questions, what strategies are used to memorize answers, how users perceive the security and memorability of their answers and how a technique which addresses key security weaknesses (but makes them less memorable) impacts users' perceptions. Our key findings reveal that despite asking participants to select security questions for an online banking scenario, participants who answered security questions with their own answers did not consider security factors. Instead, they selected easy, truthful and certain answers. Memorization strategies were ignored by most participants (even those who used unfamiliar answers). We also found that a technique designed to address key security weaknesses seemed to inspire some kind of security awareness (but would still not be enough). Based on these findings this paper provides recommendations to improve the design of security questions, strengthening fallback authentication mechanisms secure and usable.

Keywords Usable security · Fallback authentication · User behaviour · Security questions

1 Introduction

Due to their various weaknesses [1–5], security questions are considered to be neither usable (due to low memorability) nor secure enough to be used as the main account recovery mechanism [1, 3, 4, 6]. For this reason, some organizations (e.g. Google and Facebook) are moving away from using security questions and instead are using textbased and email-based password recovery [7]. This does not mean that research on security questions is no longer important [2, 5, 8]. Despite receiving a lot of scrutiny and criticism, security questions are still widely adopted [2, 3, 5, 8–10]. Hence, research still needs to investigate how the design of security questions could be improved to address their key weaknesses.

Previous security behaviour research has found that users do not understand the extent of the spectrum of the attacks that their online accounts might be vulnerable to [1, 11–13], which leads them to have misconceptions about the security threats that their accounts might be subject to. Also, despite previous research found that users' perceptions can impact the adoption of security techniques [14–18], little research [1, 2, 5, 8, 19] has investigated users'perceptions of security questions with the aim of improving their design. That is, What do users consider when selecting answers to security questions? What strategies do they use to memorize their answers? How do users use security questions? How memorable and secure are their answers to security questions? Hence, investigating users' perceptions of security questions is important because it could provide more insight on the factors that could be leading security questions to be subject to their current weaknesses. Consequently, these factors could lead to identifying recommendations that

✉ Nalin Asanka Gamagedara Arachchilage
nalin.arachchilage@auckland.ac.nz

Nicholas Micallef
n.micallef@adfa.edu.au

[1] University of New South Wales, New South Wales, Australia

[2] School of Computer Science, The University of Auckland, Auckland, New Zealand

Published online: 23 May 2021

could improve the design of security questions to address some of their key weaknesses.

This research does not only investigate users' perceptions of security questions. Previous research has continuously proposed new techniques [20–22] to address some of their key weaknesses. However, there is no evidence which shows that besides improving the key weaknesses of security questions, these techniques are also impacting users' perceptions in the required manner (e.g. by showing that users now understand the extent of the spectrum of the attacks that their accounts might be vulnerable to [1, 11–13]). Impacting users' perceptions is important because if users do not feel the need to use more secure answers (since now they understand that their accounts are vulnerable to a number of attacks), then they would still not adopt the technique that was designed to improve key security weaknesses of security questions [14–18]. Hence, the novelty of this research is that it also investigates the impact of these new techniques on users' perceptions. The aim is to understand which perceptions are actually being impacted by these new techniques (or not being impacted) so that we could provide stronger recommendations on how to further improve the design of security questions (even when using techniques that try to reduce their weaknesses).

To investigate users' perceptions, we conducted a study with 30 participants. Besides the standard security questions/answers used on online websites, in our study we also evaluate the use of system-generated information to answer security questions [22, 23]. We selected this technique because since previous research found that security questions have both security [19, 24–27] and memorability weaknesses [1, 28, 29], we wanted a technique which had the potential of impacting both security and memorability perceptions. In this way we could investigate whether new techniques could impact users' perceptions in the required manner. Hence, half of the participants (Control Group) used their own answers to security questions, while the other half (Experimental Group) used answers to security questions that were based on system-generated information (in the form of 2 system generated profiles (see Fig. 1) as proposed by [23]). In this study our participants were asked to (1) select 3 sets of security questions and answers and explain how they made their selections (same setup as [30, 31]); (2) memorize their answers and discuss the strategies used to memorize them; (3) familiarize themselves with the answers through the use of a game [32] (since the Experimental Group did not have any previous experience of using system-generated information to answer security questions); (4) recall the answers to security questions (so that Experimental Group could reflect on how easy/difficult it was to remember their answers during this study); and (5) discuss the use of security questions and their perceptions of the security and memorability of their answers (before and after the study).

Therefore, the objectives of this research were as follows: **OB1:** To understand the factors that users consider when selecting answers to security questions and whether these factors would differ when using a technique designed to address key security weaknesses. **OB2:** To understand the strategies used when memorizing answers to security questions and whether these strategies would differ when using a technique which uses less memorable answers. **OB3:** To understand users' security and memorability perceptions of their answers and whether using a technique designed to address key security weaknesses (but which uses unfamiliar answers) changes these perceptions. **OB4:** To understand the general use of security questions and whether using a technique designed to address key security weaknesses would change their use.

This work contributes to the security questions research through the following key findings: (1) When selecting and using answers to security questions, participants who answered security questions with their own answers (Control Group) did not consider security factors (despite being told that they were selecting security questions/answers for an online banking scenario). Instead, they selected answers which they considered to be easy to remember, truthful and certain; (2) The main reason why our participants did not use any memorization strategies (even those using unfamiliar answers) was that they presume that they will remember the answers to security questions (i.e. high self-efficacy); (3) Using a technique to answer security questions which focused on addressing key security limitations (through the use of system-generated information) seemed to inspire some kind of security awareness in our participants; and (4) Using a technique designed to impact both security and memorability perceptions, was not enough to change some of the misconceptions (both memorability and security related) that users have regarding security questions. Further research needs to be conducted to investigate how the used technique (i.e. system-generated information) or other techniques could be used to have a higher impact on users' perceptions.

This paper is structured in the following way. Section 2 describes the previous work related to this research. Section 3 describes the design of the system-generated profiles and security questions used in this research. Section 4 describes the methodology used to address our objectives. Section 5 presents the key findings for all our objectives **(OB1–OB4)**. In Section 6 we discuss how our key findings could improve the security and memorability weaknesses of security questions, together with the impact of using a technique which addresses key security weaknesses on users' perceptions. Section 7 provides a detailed description of the design recommendations that were extracted from our findings. In Section 8, we describe the main limitations of our research and present how we plan to address these limitations in our future work. Finally, Section 9 concludes this paper.

2 Related work

This section starts by summarizing the previous research that tried to understand user's security perceptions. Afterwards, we provide a brief overview of the latest research that investigated new techniques which have the aim of addressing key weaknesses of security questions. Finally, we conclude this section by providing a detailed overview of the work that has investigated the use of system-generated information, since this is the technique that we use in our research to investigate whether a technique that addresses key security limitations (but uses unfamiliar answers) impacts users' perceptions.

2.1 User's security perceptions

Previous research has used mental models [33, 34] to identify how non-expert users perceive security and privacy. For instance, Wash [12] studied the reasoning behind computer users' security decisions and how that affects the use of security advice. The main outcome for this research was the definition of 8 "folk models" [12, 35] of security threats that users use to justify why they ignore security advice. Subsequent research proposed 5 possible models to explain complex security concepts to non-expert users [36]. These research papers have some common findings: (1) users want security depending on context; and (2) their security needs are based on how they perceive security risk. Research has also found that non-expert users' mental models are mostly different from those of experts [34, 37]. A mental modal approach was also used by Bravo-Lillo et al. [38] to understand how users perceive and respond to computer alerts and by Ramokapane at al. [39] to understand cloud deleting practices and coping mechanisms.

In the context of authentication, Ur et al. [13] used an interview process to study password creation to investigate users' perceptions of password security. Aviv et al. [40] investigated users' perceptions of Android graphical unlock patterns by asking users to rate their perceptions of security and memorability of two unlock patterns. Most recently, Ur et al. [11] used a quantitative approach to understand whether users' perception of password security matched reality. Our work is different from previous research [11, 13, 40] since we focusing on security questions and not password or pattern authentication. Since, in our work we are also trying to understand whether users' security and memorability perceptions change when using a technique which addresses key security weaknesses (through the use of system-generated information), we used a mixed methods approach in which we focus primarily on interviews (as in [13]). However, we also use quantitative methods to back our findings.

2.2 Security questions

More recently, security questions have been investigated with the aim of determining their strengths and weaknesses [2, 3, 5, 8]. The literature highlights the following key weaknesses: (1) users tend to forget the answers to security questions after three months [28, 29], most recently Bonneau et al. [1] found that 40% of users were unable to recall their answers; (2) some answers can be guessed by searching on online public records [24] or social networking sites [19, 25]; (3) other answers can be guessed by choosing the most popular answers [27, 41]; and (4) some answers can also be guessed by close friends, family members and acquaintances [26, 42].

Recent research [2, 3, 5, 8, 20, 21, 43] has investigated different techniques to address these key weaknesses of security questions. For instance, Anvari et al. [8] proposed an inexpensive approach of generating security questions through memorizable stories for better protection of user privacy. Some of the latest research on security questions tried to leverage the use of sensors on smartphones [20, 21] to extract autobiographical information [44] about the users' smartphone behaviour during the last few days. This data is then used to define security questions about recent smartphone use [45]. Although these novel security questions techniques have managed to achieve memorability rates of about 95% using a diverse set of questions [10, 46], these techniques have the limitation that users that do not have much activity on their phones would be prone to security vulnerabilities due to a very limited answer space.

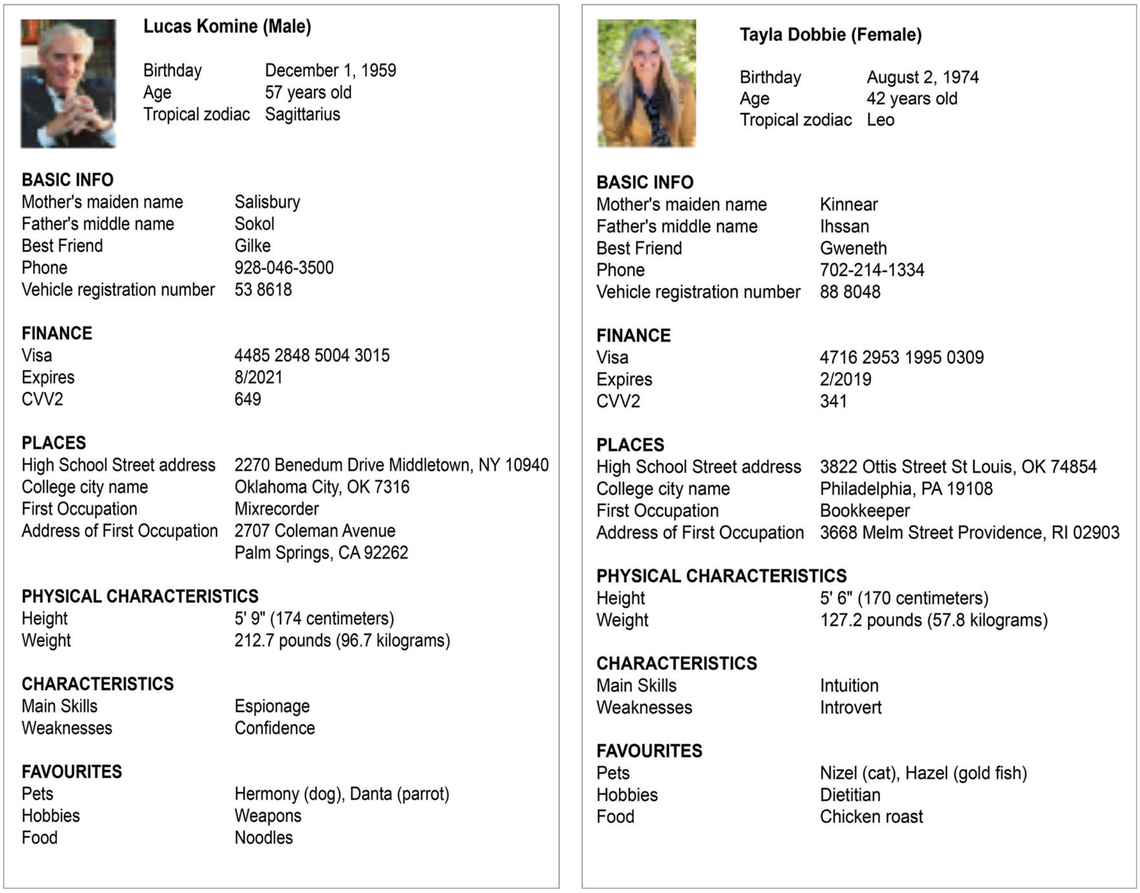
Besides the previously described work on autobiographical security questions, novel research on security questions has also investigated life-experiences passwords, in which questions are defined based on users' past experience, such as a trip, graduation, or wedding, etc. [43]. These life-experience passwords [43] were evaluated to be stronger than passwords and less guessable than security questions. However, the memorability after 6 months was still about 50%. These techniques provide an improvement over the current implementation of security questions. However, since in our research we are looking for a technique which has the potential of impacting both memorability and security perceptions we decided to review other techniques which could be more appropriate for the purpose of our research.

2.3 System-generated information

To address some of the key security vulnerabilities of security questions, previous research has proposed the use of system generated information [23, 47, 48]. Previous research found that system-generated password schemes were more secure than user-defined passwords [47, 48]. Also, previous research confirmed that system-generated password schemes were not memorable [47, 49], even when using natural-language words [47, 50]. For instance, Wright et al. [50] evaluated the usability of 3 system-generated password schemes and found that these schemes did not have sufficient memorability rates. Also, Forget et al. [51] evaluated a hybrid scheme which uses both user-selected and system-generated passwords by having a system which randomly adds characters to a user chosen password to improve its' security. This scheme only achieved a memorability of 25% when two random characters were inserted.

In an attempt to improve the usability of system-generated information, Micallef and Just [23] proposed the use of system generated information in the form of a fictitious person, and then this fictitious information would be used to answer security questions. Micallef and Just [23] reported that by using information from these fictitious profiles to answer security questions they managed to achieve a memorability of more than 33% of the security answers (for about 100 users) after 2 weeks, without providing any kind of rehearsals.

System-generated information could address some key security weaknesses of security questions by being designed to minimize guessing attacks [47, 52] by providing high entropy answers and by being less prone to social engineering attacks, since data would not be publicly available online or accessible through social networks [32]. Recent research has also investigated the design of games to help users improve the memorability of system-generated information to answer security questions [53]. This means that if these novel techniques [32, 48, 49, 53] proof to be successful in helping users remember system-generated information than this technique would have the potential to address both memorability and security weaknesses of security questions. Since answering security questions using system-generated information has the potential of impacting both memorability and security perceptions (since as outlined in the beginning of this section security questions



Lucas Komine (Male)

Birthday	December 1, 1959
Age	57 years old
Tropical zodiac	Sagittarius
BASIC INFO	
Mother's maiden name	Salisbury
Father's middle name	Sokol
Best Friend	Gilke
Phone	928-046-3500
Vehicle registration number	53 8618
FINANCE	
Visa	4485 2848 5004 3015
Expires	8/2021
CVV2	649
PLACES	
High School Street address	2270 Benedum Drive Middletown, NY 10940
College city name	Oklahoma City, OK 7316
First Occupation	Mixrecorder
Address of First Occupation	2707 Coleman Avenue Palm Springs, CA 92262
PHYSICAL CHARACTERISTICS	
Height	5' 9" (174 centimeters)
Weight	212.7 pounds (96.7 kilograms)
CHARACTERISTICS	
Main Skills	Espionage
Weaknesses	Confidence
FAVOURITES	
Pets	Hermony (dog), Danta (parrot)
Hobbies	Weapons
Food	Noodles

Tayla Dobbie (Female)

Birthday	August 2, 1974
Age	42 years old
Tropical zodiac	Leo
BASIC INFO	
Mother's maiden name	Kinnear
Father's middle name	Ihssan
Best Friend	Gweneth
Phone	702-214-1334
Vehicle registration number	88 8048
FINANCE	
Visa	4716 2953 1995 0309
Expires	2/2019
CVV2	341
PLACES	
High School Street address	3822 Ottis Street St Louis, OK 74854
College city name	Philadelphia, PA 19108
First Occupation	Bookkeeper
Address of First Occupation	3668 Melm Street Providence, RI 02903
PHYSICAL CHARACTERISTICS	
Height	5' 6" (170 centimeters)
Weight	127.2 pounds (57.8 kilograms)
CHARACTERISTICS	
Main Skills	Intuition
Weaknesses	Introvert
FAVOURITES	
Pets	Nizel (cat), Hazel (gold fish)
Hobbies	Dietitian
Food	Chicken roast

Fig. 1 System-generated profiles

Table 1 Security questions

Type	Security questions
Names	Mother's maiden name, Father's middle name, Best friends name.
Favourites	Favourite pet, Favourite food, Favourite hobby.
Numbers	Last 6 digits Visa no, Last 6 digits Phone number, Vehicle registration number.
Places	High school city name, College city name, First work city name.
Characteristics	First occupation, Last gained skill, Main Weakness.

have both memorability and security weaknesses), in our research we decided to use this technique. In the next section we discuss how we designed the system-generated information in the form of 2 system-generated profiles.

3 System-generated profiles and security questions design

The system-generated profiles (see Fig. 1) were designed to address some of the key security weaknesses that current implementation of security questions have. For instance, we verified that most attributes do not have a limited answer space. Also, since the profiles were system-generated, they are already less vulnerable towards social engineering attacks (i.e. obtaining answers from publicly available online databases or social networking websites).

Thus, to define these system-generated profiles we first identified the most popular security questions categories (i.e. names, favourites and places) that are used on websites [1, 19, 26, 30]. We then added numeric attributes (e.g. finance) because there are people that are better at remembering numbers, rather than text [54]. The characteristics categories were included to make the system-generated profiles look more real, as proposed in [23]. Similar to [23], we used Fake Name Generator[1] to define the content of the profiles. A male and female profile were selected to provide participants with the two most common genders. Since the design of system-generated profiles is not the main focus of this research, further research needs to be conducted to identify the optimal attributes that are required for a system-generated profile to be used to answer security questions.

To address the study objectives **(OB1–OB4)** a list of security questions was required so that participants could select a set of security questions and answers. Based on the system-generated profiles defined in this section we defined the security questions listed in Table 1. During the study, participants were asked to select three security questions because Renaud and Just [55] found that posing three or more questions serially would be more secure, since it is difficult to guess all three answers irrespective of how close the attacker is to the victim. Freely chosen security questions (e.g. the user can define his own security question and its answer) with free-form answers were not used in the study because previous work [31] found that there are serious concerns over the usability of these security questions (e.g. difficult to precisely remember the given answers).

4 Methodology

A study was conducted to address the main objectives of this research **(OB1–OB4)**. A between subjects design was used to understand the impact of using a technique which addresses key security weaknesses when (1) selecting security questions and answers **(OB1)**, (2) memorizing answers to security questions **(OB2)**; (3) collecting perception of the level of security and memorability of answers to security questions **(OB3)**; and (4) using answers to security questions **(OB4)**. Hence, participants were randomly split into 2 equal groups:

Control Group Participants in this group were provided with a list of security questions (see Table 1). Then, they were asked to come up with answers to those security questions by themselves.

Experimental Group Participants in this group were given 2 system-generated profiles (see Fig. 1) and they were asked to select one profile. Afterwards, they were asked to select security questions/answers from the chosen profile using the list of security questions that was provided to them (see Table 1).

Both groups followed the procedure outlined in Table 2. University ethics was obtained before conducting the study

4.1 Procedure

In the beginning (Step 0), participants were provided with an information sheet. In the information sheet participants were briefed about all the steps that they will carry out during the study. If they wanted to participate in the study,

[1] https://www.fakenamegenerator.com/

Table 2 Study procedure

Steps	Description
Step 0 — Study preparation	Study was explained to participants (through an information sheet). Participants that agreed to participate in the study were asked to sign a consent form.
Step 1 — Pre-Study interview (see Appendix 1)	Participants were interviewed about their use and perceptions of security questions.
Step 2 — Security questions selection (see Appendix 2)	Participants were provided a list of security questions (see Table 1). Participants in the Experimental Group were also provided with 2 system-generated profiles (see Fig. 1). Participants in the Control Group were asked to select 3 security questions and come up with answers. Participants in the Experimental Group were asked to select a system-generated profile (see Fig. 1) and then select 3 security questions from the provided list based on the selected profile. Participants were asked to explain the motivation behind their choices of security questions and answers.
Step 3 — Memorizing answers (see Appendix 3)	Participants were asked to memorize the answers of the security questions that they selected in Step 2, since they would be asked to recall the answers at later stage of the study. Participants were asked to discuss the strategies used to memorize the answers to security questions.
Step 4 — Familiarization task Characteristics	Participants were asked to play a game on a provided mobile device, to help them familiarize themselves with the answers of the security questions that they selected.
Step 5 — Recalling answers	Participants were asked to write down the answers to the security questions that they provided in Step 2.
Step 6 — Post-Study Interview (see Appendix 4)	Participants were interviewed about their use and perceptions of security questions that they used in the study.

participants were asked to sign a consent form. Participants were also informed that they had the option to drop out of the study, any time that they wanted, without the need of providing any reason. In Step 1, participants were interviewed about their use of security questions. They were asked to explain how they use security questions on different accounts (i.e. university email account, online banking account, loyalty cards) and whether they store (and encrypt) their answers to security questions [47, 52]. The aim was to understand how users currently use security questions **(OB4)**. In this interview participants were also asked to rate the level of security and memorability of the answers [40] to security questions that they currently use on their online accounts **(OB3)**. To gather more information about the security perceptions [30, 31] of their answers to security questions, at the end of this pre-study interview (Step 1) participants were also asked to rate how likely it is for (a) someone to guess their answers to security questions by doing an online search or by looking at their social networking accounts; (b) a family member to guess their answers; and (c) a friend to guess their answers.

In Step 2, participants were provided a list of security questions (see Table 1). Since in the country in which the study was being conducted security questions are used to recover passwords of most online banking systems, participants were told to consider that they were selecting security questions and answers for this kind of scenario. Participants in the Control Group were asked to select three security questions and come up with answers (similar setup as Just and Aspinall [30, 31]). Participants in the Experimental Group were also provided with 2 system-generated profiles (see Fig. 1). They were first asked to select a system-generated profile (see Fig. 1) and then select 3 security questions from the provided list based on the selected profile. After making their choices, participants were asked to explain how they selected their questions and answers. The aim of this step was to understand (1) the reasoning behind the choices of their security questions and answers; and (2) whether their reasoning would differ when using a technique designed to address key security weaknesses (but which uses unfamiliar answers) **(OB1)**.

In Step 3, participants were asked to memorize the answers that they selected in Step 2, since they would be asked to recall the answers at a later stage of the study. All participants memorized the answers to their security questions within 5 min After memorizing the answers to security questions, participants were asked to explain the strategies that they used to memorize their answers. The aim of this step was (1) to understand the strategies that participants use to memorize answers to security questions; and (2) to understand how these strategies would differ when using a technique which uses unfamiliar answers **(OB2)**.

Since the Experimental Group did not have any experience of using system-generated profiles to answer security questions, in our study we included Step 4, so that all participants would familiarize themselves with the answers to the security questions that they selected in Step 2. We conducted this familiarization task by asking all participants to play a game [32] related to their answers to security questions. This game, provided them with challenges about their selected answers to security questions [53]. On average participants spent 7 min playing the game. After completing this familiarization task participants were asked to recall their answers (Step 5) by writing down on a piece of paper the answers to the security questions that they selected in Step 2. We added this task so that participants in the Experimental Group would also experience remembering the answers to the security questions that were based on system-generated profiles. This task was important because in the post-study semi-structured interview (Step 6) we wanted to understand whether the use of system-generated profiles (i.e. unfamiliar answers) changed their memorability perception of answers to security questions.

Finally, in Step 6, participants were interviewed (i.e., semi-structured interview) about their use of security questions during the study. They were asked to explain whether they would use the answers to security questions that they used in the study on different accounts and whether they would store their answers **(OB4)**. In this interview participants were asked to rate the level of security and memorability of the answers to security questions that they used in the study. Afterwards, participants were also asked to rate how likely it is for (a) someone to guess the answers to the security questions used in this study by doing an online search or by looking at their social networking accounts; (b) a family member to guess the answers used in the study; and (c) a friend to guess their answers used in the study. The aim of this step was to understand whether the use of a technique which was designed to address key security weaknesses (but uses unfamiliar answers) changed users' memorability and security perceptions **(OB3)**.

4.2 Participants

This study primarily focuses on the 21–45 age group because (1) this is the age group that has the most number of online accounts (e.g. in a lot countries this age group is the one that has the most number of accounts for the most popular social networking websites, such as Facebook, Instagram, Twitter, SnapChat and Tumblr[2]); and (2) this is the age group that uses most devices to go online[3]. Hence, due to being so active online (through the numerous accounts that they use), through the use of ubiquitous devices, this age group is the one that can be considered to be most vulnerable to security attacks. Thus, 30 participants (12 females, 18 males) were recruited through word of mouth and personal connections. The mean age was 29 (22–45), med = 27.5. Most participants (20) were post-graduate students and the rest (10) were employed full-time. All participants (30/30) reported that they were experienced and confident with using security questions on online websites.

4.3 Qualitative data analysis

To identify themes from the collected qualitative information we used an adapted version [56] of the constant comparative method (CCM) approach [57], which has been used in HCI research to analyse qualitative feedback [58]. The interviewer recorded the participants' responses to the interview questions in the form of detailed notes. These notes were later coded by two researchers independently. Both researchers used the created codes to identify common themes from the collected feedback. In most instances, the themes identified/extracted by the two researchers were similar. In the few instances in which there was a disagreement, a third researcher was asked to break the tie.

4.4 Quantitative data analysis

To calculate statistical differences in our analysis we use two different statistical tests. Firstly, to compare the results from the two groups (Control Group — participants that used their own answers to security questions and Experimental Group — participants that used answers to security questions based on a system-generated profile) we always use Mann-Whitney U test because the data was not normalized (< 0.05 using Shapiro-Wilk normality test) and the samples were independent. Secondly, to compare the results of the individual groups in between phases (e.g. perception ratings from pre-study questionnaire to perception ratings from post-study questionnaire) we use Wilcoxon pairwise comparisons because the data was not normalized and the samples were paired. We assume that there were significant differences when $p < 0.05$.

5 Results

In this section we present how our findings address the main objectives **(OB1–OB4)** of our research.

5.1 Selecting security questions/answers (OB1)

To understand how users select their security questions and answers, in Step 2 we asked participants to first select security questions and answers and then explain what reasoning they used to make these selections. Participants were asked to consider that they were making their selections for an online banking system. Participants in the

[2] https://irp-cdn.multiscreensite.com/535ef142/files/uploaded/Sensis_Social_Media_Report_2016.pdf

[3] https://www.acma.gov.au/publications/2016-02/report/snapshot-aussie-teens-and-kids-online

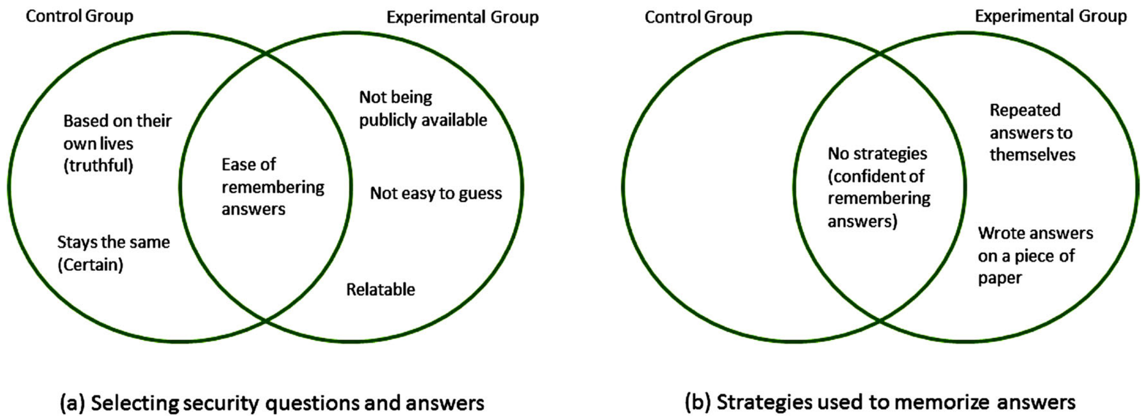

Control Group
Experimental Group
Based on their own lives (truthful)
Stays the same (Certain)
Ease of remembering answers
Not being publicly available
Not easy to guess
Relatable
(a) Selecting security questions and answers
Control Group
Experimental Group
No strategies (confident of remembering answers)
Repeated answers to themselves
Wrote answers on a piece of paper
(b) Strategies used to memorize answers

Fig. 2 **a** Main factors that affect participants selection of security questions; and **b** strategies used to memorize answers to security questions

Control Group used their own answers to security questions, while participants in the Experimental Group used system-generated profiles to answer security questions.

As shown in Fig. 2a, the main theme extracted from the feedback provided by those participants who selected their own answers to security questions (Control Group) was that they selected questions and answers that are more memorable since they would require less effort to memorize them (e.g. P1 said: “I selected security questions/answers based on memorability, so that I do not need to put any extra effort to memorize them”). Other themes extracted from these participants were that they selected answers that are based either on their own life (e.g. P3 said: “I chose the questions/answers based on their relevance to my life”) or on the fact that the answers would stay the same (i.e. favourite food might change but mother’s maiden name won’t) (e.g. P12 said: “I chose the questions because I have specific answers to them. They are more concrete, and very unlikely to change.”). These results confirm Rabkin’s [19] finding that memorability is the most important factor when selecting answers to security questions, even when asked to select security questions for an online banking scenario.

The prevailing theme extracted from participants that selected their security questions and answers based on the provided system-generated profiles (Experimental Group) was again ease of remembering answers (e.g. P10 said: “I chose the questions and answers because they are easy to remember.”). However, other popular themes (see Fig. 2a) were extracted from the feedback provided by this group of participants. The most popular being answers not being publicly available (e.g. P4 said: “Questions with answers that are not publicly available”) and answers not being prone to guessing attacks (e.g. P16 said: “I selected security questions that are not very easy to guess. Things like the “main weaknesses” of someone, which have never been exposed online. I like items which are very confidential.”). Another theme extracted from the Experimental Group participants was that they selected their answers to security questions based on relatability (e.g. P26 said: “I selected questions related to characteristics because the provided characteristics were similar to my characteristics.”). Hence, these findings indicate that even when asked to select security questions for an online banking scenario, the ease of remembering answers is the main factor considered when selecting security questions, irrespective of whether they are using familiar/unfamiliar answers (see Fig. 2a). Also, when using system-generated profiles our participants seemed to consider security as an important factor when selecting answers to security questions.

Results in Table 3 show the security questions selections for each group. These selections were categorized based on the security question types that are listed in Table 2. These results reveal that names, favourites and places were the most popular security questions, while numbers were the least popular. Hence, our results are in line with previous findings on security questions which also found these type of security questions to be the most popular amongst users [19, 26, 30]. When comparing the choices made by the different groups, one could notice that except for names and numbers the groups had opposing choices for favourites, places and characteristics. Based on the reasons that participants in the Experimental Group provided, it seems that these participants selected security questions from the favourites and characteristics categories because they found the answers provided for these categories to be more memorable, relatable and less prone to guessing attacks. Hence, the main finding from the security questions selections is that the way that system-generated profiles are designed (especially the type of answers that are provided) seem to have an impact on the types of security questions that participants select.

5.2 Memorizing answers (OB2)

After selecting the security questions/answers participants were asked to memorize their answers. The aim was to

Table 3 Choices of security questions per group

Security questions	Control group	Experimental group	Total
Names	16	14	30
Favourites	5	17	22
Numbers	3	2	5
Places	18	1	19
Characteristics	3	11	14

understand the strategies used to memorize answers to security questions and how these strategies differed when using unfamiliar answers (i.e. system-generated profiles). After giving them the required time to memorize the answers we asked participants in both groups to discuss the strategies that they used to memorize the answers. For the memorization task, we asked them to take as much time as they needed to memorize the answers, but they never took longer than 5 min to complete this task.

The main themes that were extracted from the feedback of those participants that used their own answers to security questions (Control Group) were that they do not use any memorization strategies because (1) their answers are based on themselves (e.g. P17 said: "No need to memorize,because they are based on me. I just remember them based on my life."); and (2) the answers are certain, hence, they are embedded in their memory (e.g. P7 said: "I did not make any effort because the answers are based on my past memories. They are embedded in my memory"). Hence, these findings indicate that using truthful and certain answers led our participants not to use any memorization strategies.

As shown in Fig. 2b, two prevalent themes were extracted from those participants that used system-generated profiles to select their answers to security questions (Experimental Group). They either repeated the answers to themselves for a number of times (e.g. P10 said: "Just repeated them to myself several times."), or they felt confident about answers and did not use any strategies to memorize them (e.g. P14 said: "I do not need to use any strategies because I'm confident that I will remember the answers."). Few participants also mentioned that they wrote the answers on a piece of paper (e.g. P20 said: "I memorized one of them.Then I wrote them down for a few times."). Hence, the main outcome from these findings is that our participants did not use any strategies to memorize answers to security questions (see Fig. 2b). The main reason for this seems to be that participants were confident that they would remember the answers, since they selected truthful and certain answers. With respect to those participants who used system-generated profiles, we found two contrasting views, those that came up with strategies to memorize their answers, and those who felt that they did not need to use any strategies to remember their answers (because again, they felt confident that they could remember them).

5.3 Security and memorability perceptions (OB3)

To understand users' security and memorability perceptions of their answers and whether the use of a technique designed to address key security weaknesses (but uses unfamiliar answers) changed these perceptions, we asked participants to rate the security and memorability perceptions of answers to security questions, both at the beginning (Stage 1) and at the end (Stage 6) of the study.

5.3.1 Security perceptions

In Stage 1, participants were asked to rate the level of security of the answers to security questions that they currently use on their online accounts (from 1 — not secure at all, to 5 — highly secure). The results in the second row of Table 5 show that participants seemed uncertain on whether or not their answers were secure. Afterwards, participants were asked to rate the potential of an attacker of getting access to the answers to security questions that they currently use on their online accounts, through the likelihood of "three different kind of attacks" (from 1 — not likely at all, to 5 — very likely). Participants responses to the likelihood of the attacks were based on guessing attacks from family members, friends, or from another third party who is searching for information online. When asked whether they thought that an attacker could get access to these answers by doing an online search or by looking at their social networking accounts (see third row of Table 5) participants agreed that it is quite unlikely that attackers would be able to get control of their answers in this way. These results are in line with Just and Aspinalls's findings [30, 31]. Then, we asked participants whether a family member would be capable of answering their answers to the security questions that they currently use on their online accounts. The fourth row of Table 5 shows that participants seemed to agree that a family member could guess their answers to security questions. Again, these results confirm Just and Aspinall's findings [30, 31]. Finally, we asked participants whether a friend would be capable of answering

Table 4 Security and memorability perceptions (on a 5-point Likert scale), collected in steps 1 and 6 of the study — mean (median) and standard deviation (σ)

Perceptions	Pre-study		Post-study	
	Control group	Experimental group	Control group	Experimental group
Level of security	3.3 (3) $\sigma = \pm 0.9$	3.7 (4) $\sigma = \pm 0.6$	3.1 (3) $\sigma = \pm 0.7$	3.8 (4) $\sigma = \pm 1.2$
Guessing from online search	2.6 (3) $\sigma = \pm 1.1$	2 (2) $\sigma = \pm 1$	2.9 (3) $\sigma = \pm 1$	2.1 (2) $\sigma = \pm 1.1$
Guessing from family members	3.8 (4) $\sigma = \pm 1.2$	3.4 (4) $\sigma = \pm 1$	3.8 (4) $\sigma = \pm 1.1$	2.1 (2) $\sigma = \pm 1.2$
Guessing from a friend	2.9 (3) $\sigma = \pm 1.1$	2.3 (2) $\sigma = \pm 0.6$p	3.3 (4) $\sigma = \pm 0.9$	2 (2) $\sigma = \pm 1$
Level of memorability	4.4 (4) $\sigma = \pm 0.5$	3.8 (4) $\sigma = \pm 0.8$	4.6 (5) $\sigma = \pm 0.5$	3.6 (4) $\sigma = \pm 0.8$

the answers to the security questions that they currently use on their online accounts. The fifth row of Table 5 shows that participants seemed not sure whether a friend could know their answers to security questions. These results are slightly different from Just and Aspinals's findings [30, 31]. These findings indicate that our participants seemed to believe that only family members could guess their answers to security questions, while they showed a certain level of uncertainty towards attacks that could come from people that are not so close (including friends). This level of uncertainty towards other attacks seemed to have affected their overall perception of the level of security of their answers, since they seemed to be quite unsure.

After conducting the study (Step 6), we asked all participants to rate their security perceptions based on the answers to the security questions that they used in the study. There were no changes in perceptions for the Control Group — participants who used their own answers to security questions (see second column and fourth column in Table 4). With respect to the Experimental Group there were significant changes in perception ($Z = -1.996$, $p = 0.046$) only with regard to whether a family member would be capable of answering the answers to the security questions (which were based on system-generated profiles), see third column and fifth column in Table 4. These findings indicate that the use of system-generated profiles only changed users' perception towards attacks that could come from people that are close to the participants (family members).

Table 5 Security and memorability perceptions collected in Step 1 of the study — mean (median) and standard deviation (σ)

Perceptions	Rating
Level of security	3.4 (3) $\sigma = \pm 0.8$
Guessing from online search	2.4 (2) $\sigma = \pm 1$
Guessing from family members	3.6 (4) $\sigma = \pm 1.2$
Guessing from a friend	2.7 (2) $\sigma = \pm 1$
Level of memorability	4.1 (4) $\sigma = \pm 0.7$

5.3.2 Memorability perceptions

In Step 1, participants were asked to rate the level of memorability of the answers to security questions that they currently use on their online accounts (from 1 — not memorable at all, to 5 — very memorable). The results in the fifth row of Table 5 show that participants seemed sure that the answers that they usually use are memorable. At the end of the study (Step 6), we asked all participants about the level of memorability of the answers that they used during the study. There was no significant change in the perception of the level of memorability of the answers for both groups (see fifth row in Table 4). We found significant differences, when comparing the perception of the level of memorability between the two groups ($U = 37.5$, $p = 0.001$). However, despite finding significant differences between the two groups, our findings indicate that the use of unfamiliar answers (i.e. short-term use of system-generated profiles) did not change the perception of memorability of answers to security questions because the Experimental Group participants perceived that the answers that they selected were memorable.

5.4 Using security questions (OB4)

We also wanted to understand how participants actually use security questions and answers on their online accounts and whether the use of a technique which addressed key security weaknesses would change this use.

5.4.1 On different accounts

In Stage 1 (beginning of study), we asked participants whether they use the same security questions in different accounts (i.e. university email account, online banking account, loyalty cards). Most participants (24/30) said that they do so. Our participants justified the use of the same security questions in different accounts due to (1) memorability (e.g. P8 said: "I don't want to remember various security questions"); (2) most of the websites provide the same questions (e.g. P4 said: “Most online websites have same set of questions”); and (3) convenience (e.g. P14 said “It is more convenient to use the same questions on different websites”).

Afterwards, we asked whether participants use the same answers to security questions in different accounts. Almost all participants (28/30) said that they use the same answers to security questions in different accounts. Again, most participants motivated their choice by saying that memorability is the most important factor when using security questions (e.g. P13 said: “Only if I use the same answers to security questions I can remember the answers”). The other motivations provided were (1) uniqueness of the answer (e.g. P9 said “I use same answers because I know that my answers are unique”, (2) truthfulness (e.g. P19 said: “I answer with same answers because I’m always telling the truth, the truth is the only answer”); and (3) the fact that the answers are sure and they would not change (e.g. P14 said: “I use same answers because they are the only answers that are sure”). These findings indicate that participants use the same reasoning when selecting answers to security questions on different accounts and they do not seem to vary their reasoning based on the level of security required by the account.

5.4.2 Storing answers

At the end of Stage 1, participants were also asked whether or not they store the answers to their security questions. Most participants (23/30) reported that they do not store the answers to their security questions, while 3/30 participants reported that they store them on their mobile device, 3/30 participants store the answers on their computer and 1/30 participants reported that he/she writes down their answers on a piece of paper. Most participants explained that they do not store answers because they are not bothered to do so, since security questions are only used in rare instances in which forgotten passwords need to be recovered (e.g. P29 said: “No I do not store them, since I actually never ended up in a situation in which I had to use them”). The two participants that said that they store them on a computer specifically mentioned that they store them on a draft email. All participants that store their answers (7/7) reported that they do not encrypt their answers. These findings indicate that participants confidence in their ability of remembering their answers leads them not to store them

5.4.3 Using system-generated profiles

In the post-study interview (Step 6) we asked all participants whether they would use same questions and answers on different accounts. Those participants that used their own answers to security questions (Control Group) reported that they would still use the same questions/answers in different accounts. However, half of the participants which used system-generated profiles (Experimental Group) reported that they would use different questions and answers for different accounts, if they had to use system-generated profiles to answer security questions. Also, participants that were using the system-generated profiles (Experimental Group) reported that they would like the system-generated profiles to be accessible all the time. Hence, these results confirm findings from previous research, which found that these kind of techniques increase resilience against observation attacks, but the number of users who would store their answers would increase with the complexity of the technique [47, 52]. These findings indicate that the use of system-generated profiles could change users use of security questions, since most users who used system-generated profiles changed their opinion, both regarding storing their answers and using the same answers in different accounts.

6 Discussion

In this section we discuss how the key findings from the previous section could provide more insight into the key memorability and security weaknesses that current security questions have. We also discuss the impact on users’ perceptions when using a technique designed to address key security weaknesses of security questions.

6.1 Understanding memorability weaknesses

Previous research on security questions has extensively highlighted that there is a considerable number of users that tend to forget their answers to security questions [1, 28, 29]. For instance, when analysing a large dataset of answers to security questions provided by Google, Bonneau et al. [1] found that 40% of users do not remember their answers. However, there is little empirical research which explains why users might be providing wrong answers to security questions. To address this issue, in our study, we investigated users’ reasoning when selecting and memorizing answers to security questions. Our main findings reveal that in most cases the participants that used their own answers to security questions selected answers that they considered to be easy to remember, since they selected truthful and certain answers. Our findings also show that our participants perceived that it is quite likely that they will remember the answers to their security questions (even when using unfamiliar answers). Hence, these findings indicate that our participants seemed to have quite a high level of confidence about their ability (i.e. self-efficacy) of remembering their answers to security questions, which previous research found that might not always be the case [1, 28, 29]. These findings indicate that since previous research found that answers to security questions are more likely to be used after a long time [1], users need to take it less for granted that they would remember their answers, even if they are using truthful answers that are based on certain events.

This high level of confidence about the ability (i.e. self-efficacy) of remembering their answers to security questions, seems to be one of the factors that led our participants not to use any strategies to memorize their answers. Even some of those participants that used unfamiliar answers (i.e. system generated profiles) did not feel the need to use any strategies to memorize their answers. This reluctance to memorize answers to security questions, indicates that users need to be more open to the idea of undertaking memorization and rehearsals exercises to improve the likelihood that they will remember their answers even in the long-term. One could argue that this reluctance might originate from the fact that current implementations of security questions do not encourage users to memorize their answers, and do not provide rehearsal techniques at regular intervals. Hence, the absence of these memorization-related features could be one of the factors that might be causing users to forget their answers to security questions. Our findings indicate that to address these issues, the design of security questions needs to be enhanced in such a way that it would provide users with appropriate memorization and rehearsals techniques.

6.2 Understanding security weaknesses

Previous research that investigated users' perceptions of passwords revealed that users are aware of how to improve the strength of their passwords, but they just do not do it [11]. Although our study was not structured to reveal whether our participants were aware of how to improve the security of their answers to security questions, we similarly found that the participants that used their own answers to security questions (Control Group) just did not consider security (or mentioned any word related to it) when selecting answers to security questions (despite the fact that they were briefed that they were selecting security questions/answers for an online banking system). Hence, this lack of security consideration seems to be one of the main reasons why current security questions mechanisms are prone to a number of security weaknesses [19, 24–27].

As described in the related work section, non-expert users seem to not realize the extent of the security risks that their online behaviour might be subject to [34, 37]. This might be related to the fact that users do not understand the extent of the spectrum of the attacks that their online accounts might be subject to [1, 11–13]. In fact, despite that the media has continuously highlighted numerous occurrences of celebrity account thefts (e.g. Sarah Palin's 2008 Yahoo email account hack[4] or Scarlett Johansson, Mila Kunis, and Christina Aguilera's email accounts hack[5]), our participants reported that their answers could be only guessed by family members. This might be one of the reasons why our participants did not consider any security factors when selecting answers to security questions. Our findings indicate that a possible way to address these misconceptions with security perceptions could be to enhance the design of security questions by showing the strength of the selected answers, together with showing how the selected answers could be compromised by malicious attackers and the consequences of having these accounts compromised. This functionality needs to be available to users, irrespective of whether they are answering security questions with their own answers or with some other technique (such system-generated profiles).

6.3 Impact of using the new technique

One of the objectives of this work was to understand the impact of using a technique which addresses key security weaknesses (but uses unfamiliar answers), on users' perceptions. Despite that our participants were only exposed to the new technique (use of system-generated profiles) for about 30 min, our findings still indicate that the use of this technique seemed to inspire some kind of security awareness in some of the Experimental Group participants. Most of the participants in this group: (1) considered security factors when selecting their questions/answers; (2) reported that if they had to use system-generated profiles to answer security questions they would consider using different security questions/answers in different accounts (which would lead to a more secure behaviour); and (3) recognized that the use of system-generated profiles could minimize guessing attacks from family members. Hence, this finding indicates that using this technique for a short time still managed to change some of the misconceptions that users have about their perceptions of security questions.

The use of the new technique did not have much impact on the memorability perceptions of the system-generated answers and on the perceptions of the likelihood of friends from online searches or social media. One could argue that there was no effect on the memorability perceptions because the new technique was only used for a short period of time (i.e. about 30 min). However, one should also note that participants in the Experimental Group reported that they selected answers which they considered to be

[4] http://6abc.com/archive/6398817/

[5] http://au.pcmag.com/software/23957/news/apple-no-icloud-find-my-iphone%2Dbreach-in-celeb-photo-hack

memorable and relatable. This means that the reasoning used to select the answers might have also contributed to perceiving unfamiliar answers to be memorable.

With respect to the perceptions of the likelihood of guessing attacks on answers that were based on system-generated profiles, it seems that our participants did not change their perception because they seemed convinced that even the answers to security questions that they usually use have a low likelihood of being prone to these kind of attacks. Hence, despite using a technique designed to impact both security and memorability perceptions, we found that this technique on its own was not enough to change some of the misconceptions (both memorability and security related) that users have regarding security questions. Obtaining a higher impact on users' perceptions is important because previous research found that users' perceptions can impact the adoption of security techniques [14–18]. Hence, further research is required to investigate how the used technique (i.e. system-generated information) or other techniques could be used to have a higher impact on users' perceptions.

7 Design recommendations

Based on the previous discussion, this research has the following recommendations for improving the design of current security questions mechanisms in order to reduce the impact of some of their key weaknesses:

Reduce self-efficacy of remembering answers Our findings indicate that our participants seemed to be quite confident about their ability of remembering answers. However, previous research found that security questions are more likely to be used after a long time [1]. Hence, we recommend that designers of security questions mechanisms should implement features that notify users that their answers could be used after a long time [59] and that it is important that they consider taking appropriate actions to memorize and rehearse their answers. Also, it would be beneficial to nudge users that a considerable number of people forget their answers to security questions [1]. Users could be notified of these situations before choosing the security questions or when logging into an account which uses security questions to recover passwords.

Provide memorization and rehearsal techniques Our findings indicate that our participants did not seem very keen to use strategies to memorize their answers (even those using unfamiliar answers). Since research suggests that a considerable number of users still forgets the answers to security questions [1], we recommend that designers of security questions mechanisms should implement features, that use memorization techniques [41, 60–62], to help users store information in long-term memory [63]. For example, after selecting the answers to security questions (irrespective of whether they are using familiar or unfamiliar answers), memorization instructions/exercises could be provided to help users memorize the information [64]. Since security questions are mostly used after a long time [1], we also recommend to implement some kind of rehearsals features to help strengthening the storage of information in long-term memory. The stronger the association with the information the more likely it is that users will remember the answers to their security questions [32]. This feature could be implemented by nudging users to have a look at their answers to security questions, in order to refresh their memory.

Implement mechanisms that show strength of answers Current implementations of security questions mechanisms do not show the strength of the selected answers. Hence, we recommend that designers of online authentication mechanisms should not only implement password strength mechanisms [65] to encourage users to provide stronger passwords, but should also investigate new techniques which show the strength of answers to security questions with respect to the main types of attacks that security questions could be mostly susceptible to (i.e. social engineering attacks, publicly available online data attacks, limited answer space attacks). This recommendation is important because its pointless to have the strongest possible password if that strong password could be easily reset by getting control of the account by guessing weak answers to security questions.

Use prior experiences to improve threat awareness Our findings indicate that our participants do not seem to consider the consequences that might occur when selecting weak answers to security questions because they just do not consider any security factors when selecting their answers and they only feel that their family can guess their answers. Hence, we recommend that designers of security questions mechanisms should implement features that enable reminders of previous experiences (involving cyber-attacks of online accounts, which might have affected users themselves or close friends/acquaintances). Previous research has shown that prior experiences could affect the perceived susceptibility towards a security threat [66]. This feature could be implemented either through asking questions about prior experiences [59, 67], before choosing the security questions or through nudging users with periodic notifications [68, 69] about prior experiences, when logging into an account which uses security questions to recover passwords.

Increase security threat perception According to previous security behaviour research, users' threat perception could be increased by affecting their perceived susceptibility and severity [70, 71] of an attack. Hence, we recommend that designers of security questions mechanisms should also nudge users' perception towards potential cyber-threats that could be launched on accounts of users that use weak answers to security questions [59, 67]. Thus, it is suggested that notifications [68, 69] could appear when logging into accounts, which use security questions to recover passwords. These systems could also show scenarios in which other accounts were compromised using weak answers to security questions. Furthermore, the interface could also describe the consequences that victims suffered due to these attacks. System designers should make sure that these messages are only delivered periodically using appropriate terminology, to avoid scaring and overly alarm users [72].

8 Limitations and future work

Despite using a mixed methods approach this study is mostly a qualitative inquiry — based on a sample size of 30 participants (15 participants were in the Control Group and answered security questions with their own answers, 15 participants were in the Experimental Group and used system generated profiles to answers security questions). With this sample size theoretical saturation [73] was achieved after coding the feedback collected from about 10 to 12 participants in each group. Hence, we are confident that the results presented for **OB1**, **OB2** and **OB4** represent the thoughts and considerations of a considerable segment of the inquired population (21–45 age group).

This work opens plenty of opportunities to conduct further research to improve the design of security questions to address their key weaknesses. In our future studies we plan to implement the recommendations that we provided in this research (see Section 7), with the aim of evaluating whether these recommendations actually improve the design of security questions in the long-term. Also, since we found that a technique designed to address key security weaknesses (which was only experienced for 30 min) seemed to inspire some kind of security awareness, we plan to evaluate whether using this technique (i.e. system-generated profiles) for a longer duration could have a higher impact on users' perceptions. Further research could also run studies which compare the technique that we used in our studies (i.e. system-generated profiles) to other techniques that aim to reduce key weaknesses of security questions, with the aim of understanding which technique would have the best potential to impact users' perceptions in order to change users' misconceptions regarding security questions.

9 Conclusions

This work contributes to security questions research by investigating (with a study of $n = 30$) how users select security questions, what strategies are used to memorize answers,how users perceive the security and memorability of their answers and how a technique which addresses key security weaknesses (but uses unfamiliar answers) impacts users' perceptions.

Our key findings reveal that despite asking participants to select security questions for an online banking scenario, participants who answered security questions with their own answers did not consider security factors. Instead, they selected easy, truthful and certain answers. For this reason, memorization strategies were ignored by most participants (even those who used unfamiliar answers). We also found that a technique designed to address key security weaknesses seemed to inspire some kind of security awareness. However, despite using a technique designed to impact both security and memorability perceptions, we found that this technique on its own was not enough to change some of the misconceptions (both memorability and security related) that users have regarding security questions. Obtaining a higher impact on users' perceptions is important because previous research found that users' perceptions can impact the adoption of security techniques [14–18]. Hence, further research is required to investigate how the used technique (i.e. system-generated information) or other techniques could be used to have a higher impact on users' perceptions.

Based on these key findings our research suggests that system designers should consider implementing the following recommendations: (1) provide techniques that could reduce self-efficacy of remembering answers; (2) provide memorization and rehearsal techniques to strengthen the answers memorability in long-term memory; (3) implement techniques that show the strength of the selected answers; (4) improve security awareness by providing messages about prior experiences of compromised accounts and their consequences; and (5) increase security threat perception by nudging users perceived severity and susceptibility of the kind of attacks that an account which uses weak answers to security questions could be prone to.

Although the effectiveness of these recommendations still have to be validated in a longitudinal evaluation, we still believe that these recommendations have the potential of improving the design of security questions by reducing some of their key weaknesses. Thus, this improved design could lead to (1) enhance both the usability and security level of those online services that currently use security questions to conduct password recovery; (2) inspire online services that have stopped using security questions (due to recognizing their current key weaknesses) to restart using this mechanism to recover forgotten passwords.

Appendix 1. Step 1 — Pre-study interview

The following are the questions that were asked to participants in the Pre-study interview (Step 1). In some of the questions participants were asked to select an answer from the provided choices.

1. I am experienced and confident in using security questions for online websites, in case I have forgotten passwords?
   - Strongly Disagree
   - Disagree
   - Neither Disagree nor Agree
   - Agree
   - Strongly Agree
2. Do you use the same security questions for different accounts for online websites (i.e. university email account,online banking account, loyalty cards)?.
   - Follow-up: Please explain your answer?
3. Do you usually provide the same answers to your chosen security questions given to different accounts for online websites (i.e. university email account, online banking account, loyalty cards)?
   - Follow-up: Please explain your answer?
4. How do you store the answers to your chosen security questions for online websites?
   - Follow-up if they answered yes: Do you encrypt your stored answers to your chosen security questions?
5. How secure do you think that your answers to the chosen security questions are, if you scale from 1 to 5 (1- Not secure at all to 5 - Highly secure)?
   - Not secure at all
   - Not secure
   - Not sure
   - Secure
   - Highly Secure
6. How likely is it that someone would guess your answers to your chosen security questions by doing an online search looking at your social networking accounts (e.g., facebook, linkedin), on a scale from 1 to 5 (1 - Not likely at all to 5 - very likely)?
   - Not likely at all
   - Not likely
   - Not sure
   - Likely
   - Very Likely
7. How likely is it that a family member would guess your answers to the chosen security questions, if you scale from 1 to 5 (1 - Not likely at all to 5 - very likely)?
   - Not likely at all
   - Not likely
   - Not sure
   - Likely
   - Very Likely
8. How likely it is that a friend would guess your answers to the chosen security questions, if you scale from 1 to 5 (1 - Not likely at all to 5 - very likely)?
   - Not likely at all
   - Not likely
   - Not sure
   - Likely
   - Very Likely
9. How memorable do you think that your answers to the chosen security questions are, if you scale from 1 to (1 - Not memorable at all to 5 - very memorable)?
   - Not memorable at all
   - Not memorable
   - Not sure
   - Memorable
   - Very memorable

Appendix 2. Step 2 — Security questions selections

Scenario: Imagine that you are selecting security questions and answers to recover the forgotten passwords of your online banking website. After being briefed with the previous scenario participants were provided with the following security questions:

1. Mother's Maiden Name.
2. Father's middle name.
3. Best friends name.
4. Favourite pet.
5. Favourite food.
6. Favourite hobby.
7. Last 6 digits Visa no.
8. Last 6 digits Phone number.
9. Vehicle registration number.
10. High school city name.
11. College city name.
12. First work city name.
13. First Occupation
14. Last gained skill.
15. Main Weakness.

Participants in the Control group were asked to select 3 questions and come up with the answers themselves, while participants in the Experimental Group were asked to select 3 questions and used the provided system-generated profiles to answers the questions.

When ready all participants were asked the following questions:

- How did you select these security questions/answers?

Appendix 3. Step 3 — Memorizing answers

In this step participants were asked to memorize their answers. When finished they were asked the following question:

- What strategy did you adopt to memorize your answers?

Appendix 4. Step 6 — Post-study interview

The following are the questions that were asked to participants in the Post-study interview (Step 6). In some of the questions participants were asked to select an answer from the provided choices.

1. Would you use the same security questions that you used in the study for different accounts for online websites (i.e. university email account, online banking account, loyalty cards)?

   – Follow-up: Please explain your answer?

2. Would you use the same answers to the security questions that you used in the study to different accounts (i.e. university email account, online banking account, loyalty cards)?

   – Follow-up: Please explain your answer?

   For Control Group: 3a. Would you store the answers to the security questions that you used in this study? For Intervention Group: 3b. Would you like the system generated profile to be accessible to you anytime? or do you want to see it just once?

3. How secure do you think that the answers to the security questions that you used in this study are, if you scale from 1 to 5 (1 - Not secure at all to 5 - Highly secure)?

   – Not secure at all
   – Not secure
   – Not sure
   – Secure
   – Highly Secure

4. How likely is it that someone would guess the answers to the security questions that you used in this study by doing an online search or looking at your social networking accounts (e.g., facebook, linkedin), on a scale from 1 to 5 (1 - Not likely at all to 5 - very likely)?

   – Not likely at all
   – Not likely
   – Not sure
   – Likely
   – Very Likely

5. How likely is it that a family member would guess the answers to the security questions that you used in this study, if you scale from 1 to 5 (1 - Not likely at all to 5 - very likely)?

   – Not likely at all
   – Not likely
   – Not sure
   – Likely
   – Very Likely

6. How likely it is that a friend would guess the answers to the security questions that you used in this study, if you scale from 1 to 5 (1 - Not likely at all to 5 - very likely)?

   – Not likely at all
   – Not likely
   – Not sure
   – Likely
   – Very Likely

7. How memorable do you think that the answers to the security questions that you used in this study are, if you scale from 1 to (1 - Not memorable at all to 5 - very memorable)?

   – Not memorable at all
   – Not memorable
   – Not sure
   – Memorable
   – Very memorable

References

1. Bonneau J, Bursztein E, Caron I, Jackson R, Williamson M (2015) Secrets, lies, and account recovery: Lessons from the use of personal knowledge questions at google. In: Proceedings of the 24th international conference on world wide web, pp 141–150
2. Han JK, Bi X, Kim H, Woo SS (2020) Passtag: A graphical-textual hybrid fallback authentication system. In: Proceedings of the 15th ACM Asia Conference on Computer and Communications Security, pp 60–72
3. Dhekane R (2020) Towards a usable fallback authentication mechanism
4. Schechter S, Reeder RW (2009) 1+ 1= you: measuring the comprehensibility of metaphors for configuring backup authentication. In: Proceedings of the 5th Symposium on Usable Privacy and Security, pp 1–31
5. Xu S, Chan A, Lorber MF, Chase JP (2020) Using security questions to link participants in longitudinal data collection. Prev Sci 21(2):194–202
6. Schechter S, Egelman S, Reeder RW (2009) It's not what you know, but who you know: a social approach to last-resort authentication. In: Proceedings of the sigchi conference on human factors in computing systems, pp 1983–1992
7. Stavova V, Matyas V, Just M (2016) Codes v. people: A comparative usability study of two password recovery mechanisms. In: IFIP International Conference on Information Security Theory and Practice. Springer, pp 35–50
8. Anvari A, Pan L, Zheng X (2020) Generating security questions for better protection of user privacy. Int J Comput Appl 42(4): 329–350
9. Albayram Y, Khan MMH (2016) Evaluating smartphone-based dynamic security questions for fallback authentication: a field study. Hum-Centric Comput Inf Sci 6(1):16
10. Hang A, De Luca A, Hussmann H (2015) I know what you did last week! do you? dynamic security questions for fallback authentication on smartphones. In: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, pp 1383–1392
11. Ur B, Bees J, Segreti SM, Bauer L, Christin N, Cranor LF (2016) Do users' perceptions of password security match reality? In: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, pp 3748–3760
12. Wash R (2010) Folk models of home computer security. In: Proceedings of the Sixth Symposium on Usable Privacy and Security, pp 1–16
13. Ur B, Noma F, Bees J, Segreti SM, Shay R, Bauer L, Christin N, Cranor LF (2015) "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
14. Sun S-T, Pospisil E, Muslukhov I, Dindar N, Hawkey K, Beznosov K (2011) What makes users refuse web single sign-on? an empirical investigation of openid. In: Proceedings of the Seventh Symposium on Usable Privacy and Security, pp 1–20
15. Ion I, Langheinrich M, Kumaraguru P, Čapkun S (2010) Influence of user perception, security needs, and social factors on device pairing method choices. In: Proceedings of the Sixth Symposium on Usable Privacy and Security, pp 1–13
16. De Luca A, Hang A, Von Zezschwitz E, Hussmann H (2015) I feel like i'm taking selfies all day! towards understanding biometric authentication on smartphones. In: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, pp 1411–1414
17. Bhagavatula R, Ur B, Iacovino K, Kywe SM, Cranor LF, Savvides M (2015) Biometric authentication on iphone and android: Usability, perceptions, and influences on adoption

18. Micallef N, Just M, Baillie L, Halvey M, Kayacik HG (2015) Why aren’t users using protection? investigating the usability of smartphone locking. In: Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services, pp 284–294
19. Rabkin A (2008) Personal knowledge questions for fallback authentication: Security questions in the era of facebook. In: Proceedings of the 4th symposium on Usable privacy and security, pp 13–23
20. Zhao P, Bian K, Zhao T, Song X, Li X, Ye F, Yan W et al (2016) Understanding smartphone sensor and app data for enhancing the security of secret questions. IEEE Trans Mob Comput 16(2): 552–565
21. Hang A, De Luca A, Von Zezschwitz E, Demmler M, Hussmann H (2015) Locked your phone? buy a new one? from tales of fallback authentication on smartphones to actual concepts. In: Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services, pp 295–305
22. Anvari A, Pan L, Zheng X (2020) Generating security questions for better protection of user privacy. Int J Comput Appl 42(4): 329–350
23. Micallef N, Just M (2011) Using avatars for improved authentication with challenge questions. In: Proc. of the The Fifth International Conference on Emerging Security Information, Systems and Technologies (SECURWARE 2011)
24. Gupta P, Gottipati S, Jiang J, Gao D (2013) Your love is public now: Questioning the use of personal information in authentication. In: Proceedings of the 8th ACM SIGSAC symposium on Information, computer and communications security, pp 49–60
25. Javed A, Bletgen D, Kohlar F, Dürmuth M, Schwenk J (2014) Secure fallback authentication and the trusted friend attack. In: 2014 IEEE 34th International Conference on Distributed Computing Systems Workshops (ICDCSW). IEEE, pp 22–28
26. Schechter S, Brush AJBernheim, Egelman S (2009) It’s no secret. measuring the security and reliability of authentication via “secret” questions. In: 2009 30th IEEE Symposium on Security and Privacy. IEEE, pp 375–390
27. Bonneau J, Just M, Matthews G (2010) Wha’s in a name? In: International Conference on Financial Cryptography and Data Security. Springer, pp 98–113
28. Zviran M, Haga WJ (1990) User authentication by cognitive passwords: an empirical assessment. In: Proceedings of the 5th Jerusalem Conference on Information Technology, 1990.’Next Decade in Information Technology’. IEEE, pp 137–144
29. Podd J, Bunnell J, Henderson R (1996) Cost-effective computer security: Cognitive and associative passwords. In: Proceedings Sixth Australian Conference on Computer-Human Interaction. IEEE, pp 304–305
30. Just M, Aspinall D (2009) Personal choice and challenge questions: a security and usability assessment. In: Proceedings of the 5th Symposium on Usable Privacy and Security, pp 1–11
31. Just M, Aspinall D (2010) Challenging challenge questions: an experimental analysis of authentication technologies and user behaviour. Policy Internet 2(1):99–115
32. Micallef N, Arachchilage NAG (2017) A gamified approach to improve users’ memorability of fall-back authentication. arXiv:1707.08073
33. Volkamer M, Renaud K (2013) Mental models–general introduction and review of their application to human-centred security. In: Number Theory and Cryptography. Springer, pp 255–280
34. Asgharpour F, Liu D, Camp LJ (2007) Mental models of security risks. In: International Conference on Financial Cryptography and Data Security. Springer, pp 367–377
35. Rader E, Wash R, Brooks B (2012) Stories as informal lessons about security. In: Proceedings of the Eighth Symposium on Usable Privacy and Security, pp 1–17
36. Camp LJ (2009) Mental models of privacy and security. IEEE Technol Soc Mag 28(3):37–46
37. Ion I, Reeder R, Consolvo S (2015) ... no one can hack my mind: Comparing expert and non-expert security practices. In: Eleventh Symposium On Usable Privacy and Security ({SOUPS} 2015), pp 327–346
38. Bravo-Lillo C, Cranor LF, Downs J, Komanduri S (2010) Bridging the gap in computer security warnings: A mental model approach. IEEE Secur Privacy 9(2):18–26
39. Ramokapane KM, Rashid A, Such JM (2017) I feel stupid I can’t delete...: a study of users’ cloud deletion practices and coping strategies. In: Thirteenth Symposium on Usable Privacy and Security ({SOUPS} 2017), pp 241–256
40. Aviv AJ, Fichter D (2014) Understanding visual perceptions of usability and security of android’s graphical password pattern. In: Proceedings of the 30th Annual Computer Security Applications Conference, pp 286–295
41. Denning T, Bowers K, Van Dijk M, Juels A (2011) Exploring implicit memory for painless password recovery. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp 2615–2618
42. Haga WJ, Zviran M (1991) Question-and-answer passwords: an empirical evaluation. Inf Syst 16(3):335–343
43. Woo S, Kaiser E, Artstein R, Mirkovic J (2016) Life-experience passwords (leps). In: Proceedings of the 32nd Annual Conference on Computer Security Applications, pp 113–126
44. Das S, Hayashi E, Hong JI (2013) Exploring capturable everyday memory for autobiographical authentication. In: Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing, pp 211–220
45. Albayram Y, Khan MMH (2015) Evaluating the effectiveness of using hints for autobiographical authentication: A field study. In: Eleventh Symposium On Usable Privacy and Security ({SOUPS} 2015), pp 211–224
46. Hang A, De Luca A, Smith M, Richter M, Hussmann H (2015) Where have you been? using location-based security questions for fallback authentication. In: Eleventh Symposium On Usable Privacy and Security ({SOUPS} 2015), pp 169–183
47. Shay R, Kelley PG, Komanduri S, Mazurek ML, Ur B, Vidas T, Bauer L, Christin N, Cranor LF (2012) Correct horse battery staple: Exploring the usability of system-assigned passphrases. In: Proceedings of the eighth symposium on usable privacy and security, pp 1–20
48. Al-Ameen MN, Wright M, Scielzo S (2015) Towards making random passwords memorable: leveraging users’ cognitive ability through multiple cues. In: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, pp 2315–2324
49. Al-Ameen MN, Fatema K, Wright M, Scielzo S (2015) Leveraging real-life facts to make random passwords more memorable. In: European Symposium on Research in Computer Security. Springer, pp 438–455
50. Wright N, Patrick AS, Biddle R (2012) Do you see your password? applying recognition to textual passwords. In: Proceedings of the Eighth Symposium on Usable Privacy and Security, pp 1–14
51. Forget A, Chiasson S, Van Oorschot PC, Biddle R (2008) Improving text passwords through persuasion. In: Proceedings of the 4th symposium on Usable privacy and security, pp 1–12
52. Kelley PG, Komanduri S, Mazurek ML, Shay R, Vidas T, Bauer L, Christin N, Cranor LF (2013) The impact of length and

mathematical operators on the usability and security of system-assigned one-time pins. In: International Conference on Financial Cryptography and Data Security. Springer, pp 34–51

53. Micallef N, Arachchilage NAG (2017) Involving users in the design of a serious game for security questions education. arXiv:1710.03888
54. Milikowski M, Elshout JJ (1995) What makes a number easy to remember? Br J Psychol 86(4):537–547
55. Renaud K, Just M (2010) Pictures or questions? examining user responses to association-based authentication. Proceedings of HCI 2010 24, pp 98–107
56. Baillie L (2002) The home workshop: A method for investigating the home. Ph.D. Thesis, Edinburgh Napier University
57. Glaser BG, Strauss AL, Strutzel E (1968) The discovery of grounded theory; strategies for qualitative research. Nurs Res 17(4):364
58. Micallef N, Baillie L, Uzor S (2016) Time to exercise! an aide-memoire stroke app for post-stroke arm rehabilitation. In: Proceedings of the 18th international conference on Human-computer interaction with mobile devices and services, pp 112–123
59. Vance A, Eargle D, Ouimet K, Straub D (2013) Enhancing password security through interactive fear appeals: A web-based field experiment. In: 2013 46th Hawaii International Conference on System Sciences, pp 2988–2997, IEEE
60. Stobert E, Biddle R (2013) Memory retrieval and graphical passwords. In: Proceedings of the ninth symposium on usable privacy and security, pp 1–14
61. Castelluccia C, Dürmuth M, Golla M, Deniz F (2017) Towards implicit visual memory-based authentication
62. Stobert E, Biddle R (2014) A password manager that doesn't remember passwords. In: Proceedings of the 2014 New Security Paradigms Workshop, pp 39–52
63. Atkinson RC, Shiffrin RM (1968) Human memory: A proposed system and its control processes
64. Juang KA, Ranganayakulu S, Greenstein JS (2012) Using system-generated mnemonics to improve the usability and security of password authentication. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, vol. 56, pp 506–510, SAGE Publications Sage
65. Komanduri S, Shay R, Kelley PG, Mazurek ML, Bauer L, Christin N, Cranor LF, Egelman S (2011) Of passwords and people: measuring the effect of password-composition policies. In: Proceedings of the sigchi conference on human factors in computing systems, pp 2595–2604
66. Mwagwabi F, McGill T, Dixon M (2014) Improving compliance with password guidelines: How user perceptions of passwords and security threats affect compliance with guidelines. In: 2014 47th Hawaii International Conference on System Sciences. IEEE, pp 3188–3197
67. Briggs P, Jeske D, Coventry L (2017) Behavior change interventions for cybersecurity. In: Behavior change research and theory, pp 115–136, Elsevier
68. Felt AP, Ainslie A, Reeder RW, Consolvo S, Thyagaraja S, Bettes A, Harris H, Grimes J (2015) Improving ssl warnings: Comprehension and adherence. In: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, pp 2893–2902
69. Micallef N, Just M, Baillie L, Alharby M (2017) Stop annoying me! an empirical investigation of the usability of app privacy notifications. In: Proceedings of the 29th Australian Conference on Computer-Human Interaction, pp 371–375
70. Liang H, Xue Y (2010) Understanding security behaviors in personal computer usage: A threat avoidance perspective. J Assoc Inf Syst, 11(7):394–413
71. Tsai H-YS, Jiang M, Alhabash S, LaRose R, Rifon NJ, Cotten SR (2016) Understanding online safety behaviors: A protection motivation theory perspective. Comput Secur 59: 138–150,
72. Sasse A (2015) Scaring and bullying people into security won't work. IEEE Secur Privacy 13(3):80–83
73. Bowen GA (2008) Naturalistic inquiry and the saturation concept: a research note. Qual Res 8(1):137–152