Towards Making Random Passwords Memorable: Leveraging Users' Cognitive Ability Through Multiple Cues

Mahdi Nasrullah Al-Ameen
Department of CSE
The University of Texas at
Arlington
Arlington, TX, USA
mahdi.al-
ameen@mavs.uta.edu

Matthew Wright
Department of CSE
The University of Texas at
Arlington
Arlington, TX, USA
mwright@uta.edu

Shannon Scielzo
Department of Psychology
The University of Texas at
Arlington
Arlington, TX, USA
scielzo@uta.edu

ABSTRACT

Given the choice, users produce passwords reflecting common strategies and patterns that ease recall but offer uncertain and often weak security. System-assigned passwords provide measurable security but suffer from poor memorability. To address this usability-security tension, we argue that systems should assign random passwords but also help with memorization and recall. We investigate the feasibility of this approach with *CuedR*, a novel *cued-recognition* authentication scheme that provides users with multiple cues (visual, verbal, and spatial) and lets them choose the cues that best fit their learning process for later recognition of system-assigned keywords. In our lab study, all 37 of our participants could log in within three attempts one week after registration (mean login time: 38.0 seconds). A pilot study on using multiple CuedR passwords also showed 100% recall within three attempts. Based on our results, we suggest appropriate applications for CuedR, such as financial and e-commerce accounts.

Author Keywords

Usable security; authentication; cued-recognition

ACM Classification Keywords

K.6.5 Management of Computing and Information Systems: Security and Protection- Authentication

INTRODUCTION

In most systems, users are tasked with creating a password that should be both secure and memorable. Users, however, typically lack information about what is secure in the face of modern cracking and attacks tools, as well as how to construct memorable strings, memorize them quickly, and accurately recall them later. Faced with this challenge, users often create passwords that may seem secure and memorable but fail on one or both counts. Failure to understand security requirements leads to guessable passwords, while memorability

issues lead not only to inconvenience, but also to password reset systems that are often abused by hackers [8, 36].

We argue that the burden of password creation should be borne by the system, rather than the user. With system-assigned passwords, the user does not have to guess whether a password is secure, and the system can ensure that all passwords offer the desired level of security. Additionally, while password reuse could pose a serious security threat [14], using system-assigned passwords ensures that users do not reuse a password (or modification thereof) already used on another account.

Making system-assigned passwords memorable, however, has proved challenging. Different variants of system-assigned passwords have been proposed [18, 21, 38, 49], but none of them provides sufficient memorability. We postulate that new authentication systems should more effectively make use of humans' cognitive strengths and accommodate users with different learning styles. To this end, we draw upon several prominent theories of memory to design *CuedR*, a novel authentication scheme that offers visual, verbal, and spatial cues to help users recognize system-assigned keywords.

Contributions

Memorability. In CuedR, the system assigns users six keywords, each from a distinct portfolio (e.g., animals, fruits, or vehicles) of 26 keywords. Both at registration and at login, users are provided with an image of the keyword (a visual cue); a number and a phrase associated with the keyword (verbal cues); and the fixed position of all of the elements on the page (spatial cues). Users with different learning styles can focus on the cues that help them best remember the keyword. Moreover, the cues facilitate an elaborative encoding that helps to transfer the keywords from the working memory to long term memory at registration [4], helping users recognize their keywords when logging in later.

In our single-password study, all 37 participants remembered their CuedR password after one week of registration. We note that no other system-assigned password scheme has reported 100% memorability to our knowledge, even schemes offering only PIN-replacement security levels (e.g. 13 bits of entropy). Despite high login times (38.0 seconds on average), participants reported high levels of satisfaction with the scheme and 84% preferred to use it in real life as a replacement to traditional textual passwords.

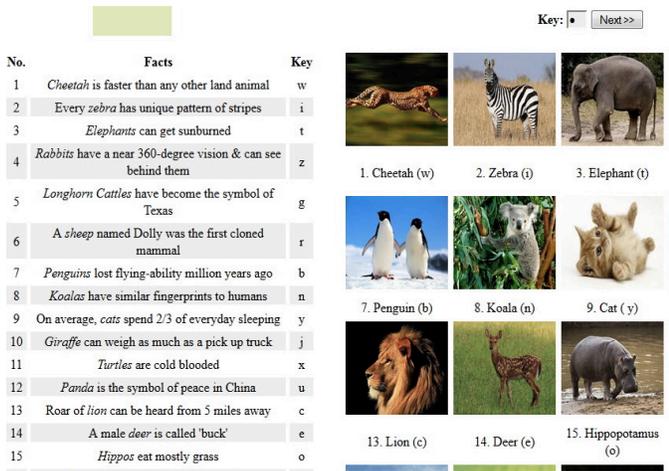| No.	Facts	Key
1	*Cheetah* is faster than any other land animal	w
2	Every *zebra* has unique pattern of stripes	i
3	*Elephants* can get sunburned	t
4	*Rabbits* have a near 360-degree vision & can see behind them	z
5	*Longhorn Cattles* have become the symbol of Texas	g
6	A *sheep* named Dolly was the first cloned mammal	r
7	*Penguins* lost flying-ability million years ago	b
8	*Koalas* have similar fingerprints to humans	n
9	On average, *cats* spend 2/3 of everyday sleeping	y
10	*Giraffe* can weigh as much as a pick up truck	j
11	*Turtles* are cold blooded	x
12	*Panda* is the symbol of peace in China	u
13	Roar of *lion* can be heard from 5 miles away	c
14	A male *deer* is called 'buck'	e
15	*Hippos* eat mostly grass	o

Figure 1. A partial screen shot during login. The facts corresponding to each keyword appear on the left side of the screen. The key is shown in parenthesis next to each keyword and also in the rightmost column of the table

Security. By using system-assigned random passwords, the effective entropy of the passwords is equal to the theoretical entropy, which is set to 28 bits in our studies. Additionally, CuedR provides *variant response* during login, which is known to be an important feature to gain resilience against observation attacks (e.g., shoulder surfing, keystroke loggers) [5].

RELATED WORK

In this section, we give a brief overview of notable textual and graphical password schemes, in which we highlight why existing schemes are insufficient.

Textual Password Schemes

User-chosen passwords. Traditional user-chosen textual passwords are fraught with security problems and are especially prone to password reuse and predictable patterns [14, 40]. Das et al. [14] found that 43% of users use the identical password in multiple sites, while 30% of non-identical passwords could be cracked in less than 100 attempts. Shay et al. [40] report that password restriction policies do not necessarily lead to more secure passwords but can adversely affect memorability.

System-assigned passwords. System-assigned random textual password schemes are more secure but fail to provide sufficient memorability, even when natural-language words are used [38, 49]. Wright et al. [49] compared the usability of three different system-assigned textual password schemes: Word Recall, Word Recognition, and Letter Recall. None of these schemes had sufficient memorability rates. Forget et al. [20, 21] proposed the Persuasive Text Passwords (PTP) scheme as a hybrid between user-selected and system-assigned passwords, in which the user first creates a password and PTP improves its security by placing randomly-chosen characters at random positions in the password. Unfortunately, the memorability for PTP is just 25% when two random characters are inserted [20].

Graphical Password Schemes

Graphical password schemes can be divided into three categories [5], based on the kind of memory leveraged by the systems: i) Drawmetric (recall-based), ii) Locimetric (cued-recall-based), and iii) Cognometric (recognition-based). *Passfaces* [1], a cognometric graphical password scheme, is commercially available and deployed by a number of organizations, including banks and government agencies.[1]

Drawmetric. The user is asked to reproduce a drawing in this category of graphical passwords. In *Draw-a-Secret (DAS)* [25], a user draws on top of a grid, and the password is represented as the sequence of grid squares. Nali and Thorpe [28] have shown that users choose predictable patterns in DAS. *BDAS* [17] intends to reduce the amount of symmetry in the user's drawing by adding background images, but this may introduce other predictable behaviors such as targeting similar areas of the images or image-specific patterns [5]. DAS and BDAS have recall rates of no higher than 80%.

Locimetric. The password schemes in this category, including *Passpoints* and *Cued Click-Points (CCP)*, present users with an image and have users select points on the image as their password. Dirik et al. [16] developed a model that can predict 70-80% of users' click positions in Passpoints. To address this issue, Chiasson et al. proposed *Persuasive Cued Click-Points (PCCP)* [11, 22], in which a randomly-positioned viewport is shown on top of the image during password creation, and users select their click-point within this viewport. The memorability for PCCP was found to be 83-94%. In a follow-up study, Chiasson et al. [10] found predictability in users' click points and indicate that predictability is still a security concern for PCCP.

Cognometric. In this recognition-based category of graphical passwords, the user is asked to recognize and identify their password images from a set of distractor images. *Passfaces* [1] is a commercial cognometric system in which users select one face among a panel of nine distractor faces and repeat this over several panels. Davis et al. [15] have found that users select predictable faces, biased by race, gender, and attractiveness of faces. As a result, the commercial Passfaces [1] product now assigns a random set of faces instead of allowing users to choose. However, Everitt et al. [18] show that users have difficulty in remembering system-assigned Passfaces. Hlywa et al. [24] found no significant difference in memorability between cognometric schemes providing either face images or object images (entropy: 28 bits), while the mean login time was 31 seconds for object recognition and 41 seconds for face recognition.

In sum, schemes with lower risk of predictability also show lower recall rates. Password managers [12] fail to provide a suitable solution in this case, as it suffers from usability (in implementation) and security (e.g., single point of failure) problems. Indeed, two recent papers extensively examine security problems in a range of password managers [27, 41].

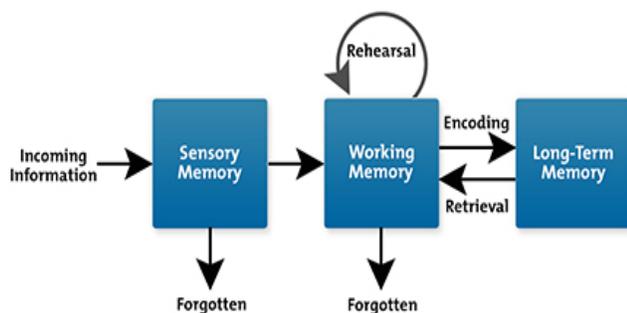[1] **http://www.realuser.com/** shows testimonials about Passfaces from customers.

Figure 2. Illustration of cognitive memory model

Thus, despite a large body of research, it remains a critical challenge to build an authentication system that offers both high memorability and guessing resilience.

CuedR: SYSTEM DESIGN

In CuedR, six keywords are randomly assigned to the user each from a distinct portfolio (e.g., animals, fruits, or vehicles), where each portfolio presents 26 keywords. To aid memorability, our scheme offers graphical, verbal, and spatial cues corresponding to each keyword. In particular, each keyword (e.g., "Zebra") has an image (a picture of a zebra), a number and a phrase related to the keyword ("2. Each zebra has a unique pattern of stripes."), and the position of both the image and phrase are fixed in both absolute terms and relative to the other images and phrases (Zebra is at the top between Cheetah and Elephant). See Figure 1 for a screenshot illustrating these features.

Each time a portfolio is loaded, each of the 26 lowercase letters a-z is assigned randomly as a *key* to one keyword on the page. The user inputs the key letter corresponding to her keyword into a single-character password field to move on to the next portfolio. The key letter changes every time to provide the *variant response* property [5]. Schemes with this property have been shown to provide higher resilience to shoulder surfing and simple keystroke loggers than schemes like traditional textual passwords in which the same letters are entered at every login [5, 43].

User authentication.   At user registration in CuedR, the system randomly selects six portfolios (without replacement) and the user is assigned one keyword from each of these portfolios. The user enters the key corresponding to the assigned keyword to get forwarded to the next portfolio. In case of a wrong entry she will immediately be informed about the error and will need to enter the correct one. During login, the user recognizes her system-assigned keyword from the portfolio and enters the key corresponding to that keyword into a small password field. A successful authentication requires the user to correctly enter keys for all six of her assigned keywords.

When the user makes a mistake during login, CuedR shows the user a portfolio that is different from her next assigned portfolio. A legitimate user can recognize this *implicit feedback* as an indicator that something was wrong and that she should go back and correct the mistake, while an attacker will

not know which portfolios are correct. Implicit feedback is thus a desirable feature to enhance usability when passwords have multiple parts [5].

Password storage.   For secure storage of the user's authentication secret, the six keywords can be concatenated together with a salt and hashed using a slow hash function like bcrypt [33] or PBKDF2 [26]. Implicit feedback can be implemented by making the selection of the next portfolio a function of the current portfolio and the keyword selected, which is independent of the correctness of the responses. Thus, correctness only needs to be checked after all keys have been entered.

THE SCIENCE BEHIND CuedR

In this section, we explain from the perspective of cognitive psychology how our design choices are set up to provide high memorability.

Long-Term Memory

We incorporate the scientific understanding of long-term memory to advance the scheme's usability properties. According to the cognitive memory model proposed by Atkinson and Shiffrin [4], any new information is transferred to short-term memory (STM) through the sensory organs, where STM holds the information as *memory codes*, or mental representations of selected parts of the information. The information is transferred from STM to long-term memory (LTM), but only if it can be further processed and encoded. In this respect, an elaborative encoding would take place if the information can be associated with something meaningful, such as *cues*. This encoding helps people to remember and retrieve the processed information efficiently over an extended period of time (see the illustration in Figure 2).

In CuedR, users focus their attention to learn keywords through associating them with the corresponding cues, which should help to process and encode the keywords in memory and store them in the LTM. The cues would assist the user to recognize the keywords in the future, which should enhance their memorability.

Memory Retrieval

We designed CuedR to require users to perform a recognition task. Researchers in psychology have found that recognition (identifying the correct item among a set of distractors) is easier than recall (reproducing the item from memory) [46] and have developed two main theories to explain this: *Generate-recognize theory* [3] and *Strength theory* [47].

Generate-recognize theory [3] speculates that recall is a two-phase process. In the generate phase, a list of candidate words is formed by searching long-term memory. Then, in the recognize phase, the list of words is evaluated to see if they can be recognized as the sought-out memory. According to this theory, recognition tasks do not utilize the generation phase and are thus faster and easier to perform. Strength theory [47] states that although recall and recognition involve the same memory task, recognition requires a lower threshold of strength that makes it easier. The point is commonly illustrated in examples from everyday life. For example, multiple

choice questions are frequently easier than essay questions since the correct answer is available for recognition.

Memory Cues

Psychology research [3, 46] has shown that it is difficult to remember information spontaneously without memory cues, and this suggests that authentication schemes should provide users with cues to aid memory retrieval. *Encoding specificity theory* [45] postulates that the most effective cues are those that are present at the time of remembering. In CuedR, cues are provided during registration, i.e., the learning period, and also at login.

Why use multiple cues?

In CuedR, the user has five different references (cues) that she can leverage to learn the keyword: (i) the image, (ii) the number from 1 to 26, (iii) the phrase or fact associated with the keyword, (iv) the absolute positions of the keyword, the image, and the phrase/fact, and (v) the positions of the keyword, image and phrase/fact relative to the other keywords, images and phrases. This combination brings together graphical (images), spatial (positions), and verbal (facts, numbers) information. Thus, a user may focus on just those cues that she finds most appropriate to her learning process, while the other cues may provide additional support for memorability.

Graphical cues. Psychology research [29, 31] reveals that the human brain is better at memorizing graphical information as compared to textual information. This is known as the *picture superiority effect*, which motivates us to include graphical cues (images) in our scheme. Several explanations for this effect have been proposed. The most widely accepted is *dual-coding theory* [31], which postulates that in human memory, images are encoded not only visually and remembered as images, but they are also translated into a verbal form (as in a description) and remembered semantically. Another is the *sensory-semantic model* [29], which states that the images are accompanied by more distinct sensory codes that allow them to be more easily accessed.

Verbal and spatial cues. While images are generally effective cues, not all users may have a strong visual memory. Additionally, many graphical password schemes require good vision and motor skills, which elderly users [34] may lack. Thus, we provide verbal and spatial cues in addition to graphical cues to let users leverage their cognitive ability in memorizing the keywords. Yan et. al. [50] examined the influence of phrases in increasing the memorability of passwords, which inspires us to accommodate a common phrase or fact for each keyword as a verbal cue.

Having a fixed set of objects in a certain place aids to augment *semantic priming*, which refers to recognizing an object through its relationship with other objects around it [1]. Semantic priming thus eases the recognition task [1]. In CuedR, the keywords and cues in a portfolio remain same and presented at a fixed position whenever that portfolio is loaded, which establishes a relationship between them and reinforces semantic priming. A recent study [42] also shows that keeping objects in a fixed position improves the usability during recognition.

CuedR: THROUGH THE LENS OF PASSWORD LITERATURE

Through a comprehensive survey on 25 different graphical password schemes, Biddle et al. [5] identified seven features that should be offered by an ideal graphical password system. The authors [5] state, "We expect tomorrow's ideal graphical password systems may have many of the following desirable characteristics, reflecting lessons learned from proposals to date."

In this section, we analyze how CuedR addresses these seven features and explain our design choices based on the findings from the literature on passwords.

[1] Theoretical password space meeting the security policy of the intended domain

Well-known recognition-based schemes, such as Passfaces [1] and Story [15], originally provided no more than 13 bits of theoretical entropy. Later, Hlywa et al. [24] conducted a study on recognition based graphical passwords with 20 bits of entropy, since 20 bits of entropy with reasonable lockout rules is considered sufficient to prevent online brute-force attacks [19]. In CuedR, we use more than 20 bits of entropy, in particular 28 bits, to maintain comparability with prior studies on system-assigned passwords [49]. During login in CuedR, a user has to recognize her keyword from a portfolio of 26 distinct keywords. This is required six times, once for each portfolio. The password space is thus $log_2(26)^6 \approx 28$ bits. CuedR can be used with a range of entropy values by varying either the number of keywords or portfolios.

[2] Avoiding exploitable reductions in security due to user choice of passwords

Statistical password distributions are often not equiprobable due to scheme-dependent predictability of user choices [6]. In CuedR, passwords are randomly assigned by the system, which provides two security benefits. First, the effective password space in CuedR is same as the theoretical space. Second, the system gains user-choice resilience and thus provides robustness against online guessing attacks that exploit password reuse, personal information and predictable strategies [14, 40].

[3] At least mild resistance to shoulder surfing and key logging, through variant response

Variant response refers to varying how the password is entered across different login sessions, which is an important feature to offer robustness against shoulder surfing and keystroke loggers [5].

Shoulder surfing

It is difficult in practice to observe both keyboard and monitor at the same time. Thus, graphical password schemes that include the variant response feature with keyboard entry provide higher resilience to shoulder surfing compared to traditional textual passwords and graphical passwords with mouse input [43]. In a study by Tari et al. [43], participants playing the role of shoulder surfers were able to gain 73% of non-dictionary passwords, 26% of dictionary passwords, and

62% of graphical passwords with mouse input, but just 11% of graphical passwords with variant response.

CuedR offers the variant response feature, where the user enters a key corresponding to her keyword using the keyboard, and watching only keyboard entries is not sufficient for a shoulder surfer, as the key associated with each keyword changes with every login attempt. The entered key is shown as an asterisk or dot (as with a regular password) to minimize the risk of shoulder surfing.

Variant response does not protect against an attacker who can use a video camera to record both the monitor and keystrokes at the same time, and attackers may gain the user's credentials when they are assigned during registration. Thus, we only claim that CuedR provides mild resistance to shoulder surfing through variant response, which conforms to the desired level of security in this regard [5]. We recommend that users register in a secure environment (e.g., avoiding public terminals) to ensure better security against shoulder surfing.

Keystroke and mouse loggers
Keystroke loggers record keyboard input and mouse loggers capture mouse actions to make the user's credentials available for retrieval by remote attackers [5]. Biddle et al. [5] state that a system provides resilience against keystroke/mouse loggers when the keyboard/mouse entries for authentication vary across subsequent login sessions. Thus, the variant response feature in CuedR offers better resilience against basic keystroke loggers compared to a password system where the same letters are entered during every login session. CuedR is clearly resilient to mouse loggers, as it does not use mouse input.

[4] Cues aiding memorability
While different variants of system-assigned passwords failed to provide satisfactory memorability [20, 21, 38, 49], CuedR achieves a good memorability through associating each keyword with a set of cues and letting users choose the appropriate one(s) to their learning process [2]. We describe the basis for the effectiveness of cues in the previous section, and we report on user perceptions of different cues in the results section.

[5] Usability as close as possible to, or better than, textual passwords
Shay et al. [39] performed a comprehensive study on the memorability of user-chosen textual passwords following different composition policies, where *basic12* was the simplest form of passwords in which the user had to create a password of at least 12 characters without any composition requirements or dictionary check. Participants reported the least difficulty to create and remember a *basic12* password. After two days, 86% of participants who wrote down their password could log in (76% on the first attempt), while 75% of participants who did not write down their password could log in (61% on the first attempt). For CuedR, after one week, we found that 100% of participants could log in (89% on the first

---

²We note that direct comparison between different studies should be taken with caution

---

attempt). So, we see that memorability is better than textual passwords with moderate security requirements.

Although the login time for CuedR is high compared with traditional textual passwords, users mostly disagreed with the notion that the scheme is too time consuming (see Table 1). Overall, users reported satisfaction with the usability of CuedR and 84% preferred to use it in real life as a replacement to traditional textual passwords.

[6] Implicit feedback to legitimate users, when passwords are multi-part
Implicit feedback instantly notifies a user when she makes a mistake, instead of showing her an error message at the end of all entries. Due to its implicit nature, this feedback should only be recognizable and useful to the legitimate user. An attacker who does not know about a user's portfolios must make all six guesses in CuedR to learn whether he has succeeded or not. Implicit feedback has already been shown to have satisfactory user acceptance in a cued-recall based scheme [13]. To accommodate this feature in CuedR, we build distinct portfolios of images (i.e., "animals", "fruits", "flowers", etc.) so that a user can clearly distinguish among the portfolios at a glance and quickly realize her mistake.

[7] Leveraging pre-existing user-specific knowledge where possible
Leveraging pre-existing user-specific knowledge, for example answering cognitive questions or recognizing personal images from decoys could make the scheme vulnerable to targeted guessing attacks (e.g., guessing by acquaintances). So, we did not include this feature in CuedR to ensure security. Since CuedR offers good memorability, it is not clear if user-specific knowledge is required, though it could help to reduce the cognitive burden on users. Exploring ways to securely leverage user-specific knowledge for authentication could be an interesting venue for future work.

USER STUDY
We now present the design of our user study to evaluate the usability and memorability of CuedR. The study procedures were approved by our university's Institutional Review Board (IRB) for human subjects research.

Participants, Apparatus and Environment
For this experiment, we recruited 37 students (25 women, 12 men) through our university's Psychology Research Pool. Participants came from diverse backgrounds, including majors from Nursing, Psychology, Business, Political Science, Biology, Physical Science, and Social Work. The age of the participants varied between 17 to 30 with a mean age of 21. They make regular use of the Internet and websites that require authentication. Each participant was compensated with course credit for participation and was aware that her performance or feedback in this study would not affect the amount of compensation.

The lab studies were conducted with one participant at a time to allow the researchers to observe the user's interaction with the system. For this study, we built 18 different portfolios

(e.g., animals, fruits, flowers, and vehicles), and collected the images (graphical cues) and phrases/facts (verbal cues) from free online resources.

Procedure
We conducted the experiment in two sessions, each lasting around 30 minutes. The second session took place one week after the first one to test memorization of the password. Note that the one-week delay is larger than the maximum average interval for a user between her subsequent logins to any of her important accounts [23]. One week is also a common interval used in authentication studies (e.g., [17, 30, 49]).

Session 1. After signing a consent form, the participants performed a practice trial with CuedR to compensate for novelty effect. We did not collect data for this practice trial. At registration, six portfolios were randomly chosen by the system and a user was assigned at most one keyword from each portfolio. Then participants were asked to spend 60 seconds in completing a mental rotation test (MRT) puzzle shown on the computer screen, which helps to clear their working memory [32]. Participants were then given questionnaire that gathered demographic information, and were asked to log into the same site with CuedR (*login 1*). They were asked to not write down their authentication secrets.

Session 2. The participants returned after one week of registration, and logged into the site using CuedR (*login 2*). After they had finished, we conducted an anonymous paper-based survey. Participants were then compensated and thanked for their time.

Ecological Validity
Our participants were young and university educated, which represents a large number of frequent Web users, but may not generalize to the entire population. They came from diverse majors including Nursing, Psychology, Physical Science, Business, etc. As the study was performed in a lab setting, we were only able to gather data from 37 participants. We believe that 37 provides a suitable sample size for a lab study as compared to the prior studies on password memorability [11, 13, 44, 48].

RESULTS
In this section, we discuss the results of our user study. We label the login performance of participants in session 1 and session 2 as *login 1* and *login 2*, respectively. We evaluated the usability of CuedR via all metrics suggested in the literature [35]: memorability, login time, number of login attempts, and user feedback. In addition, we analyzed the impact of portfolios on login performance and user perceptions on the effectiveness of different cues. We also discuss the results of pilot study on the memorability of multiple CuedR passwords.

Memorability
We observed a 100% login success rate for CuedR in both *login 1* and *login 2*. In *login 1*, all the participants successfully recognized the keywords on the first attempt. In *login 2*, 89% of participants succeeded on the first attempt to recognize all
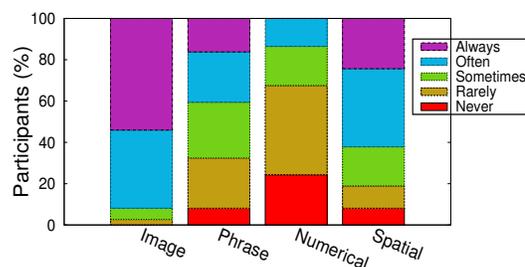

Figure 3. Responses to the question: "How often did the following cues assist you in recognizing keywords in CuedR?"

six keywords. The other four participants (11%) recognized five of out of six keywords on the first attempt. Three participants corrected their mistake on the second attempt, and the other participant succeeded on the third attempt.

Registration and Login Time
The mean time for registration was 31.2 seconds (median: 30 seconds, SD: 10.5 seconds). The mean time for successful login were 25.7 seconds (median: 24.0 seconds, SD: 8.3 seconds) in *login 1*, and 38.0 seconds (median: 39.0 seconds, SD: 11.4 seconds) in *login 2*. A paired-samples t-test reveals that login time in *login 1* was significantly less than that in *login 2*, $t(36) = 7.81$, $p < 0.01$. This was expected, as participants performed *login 1* shortly after learning the keywords. To note, the reported registration and login time include the time to download images.

The login time in CuedR is in line with that in prior recognition based schemes offering 28 bits of entropy [24, 49]. We note that our results for login time are likely conservative, since they measure initial use. A recent field study [2] reveals that login time decreases with the frequent use of a scheme due to training effects. These findings are in agreement with our user feedback, where the participants reported that with practice, they could quickly recognize the keywords (see Table 1).

Impact of Portfolios on Usability
In our study, all the participants succeeded to recognize their keywords irrespective of the type of portfolios in both *login 1* and *login 2*. In *login 1*, no participant made any mistake in any portfolio, and thus there was no difference among portfolios for the number of attempts to succeed. In *login 2*, four participants (11%) required multiple attempts to succeed (see the results for *Memorability*), where one-way ANOVA test results show that there was no significant difference among portfolios in terms of the number of attempts required to successfully recognize the keywords, $F(17, 220) = 1.16$, $p = 0.31$. In addition, we conducted a post-hoc pairwise comparison using Tukey's HSD (Honestly Significant Difference), which reveals no significant difference between any pair of portfolios for the number of attempts to succeed.

Our one-way ANOVA test results demonstrate that there was no significant difference among different portfolios in terms of the time to learn the keyword during registration, $F(17, 220) = 0.76$, $p = 0.71$, or recognize the keyword either in *login 1*, $F(17, 220) = 1.16$, $p = 0.31$, or in *login 2*,

Table 1. Questionnaire responses for the usability of CuedR. Scores are out of 10. * indicates that scale was reversed. SD: Standard Deviation

Questions	Mode	Median	Mean	SD
I could easily sign up with CuedR	10	9.0	9.0	1.3
The login using CuedR was easy	10	10.0	9.5	0.7
Keywords are easy to remember in CuedR	10	10.0	9.4	0.8
*I found CuedR too time-consuming (i.e., I did not find CuedR too time consuming)	10	7.0	6.4	2.6
With practice, I could quickly enter my password in CuedR	10	10.0	9.8	0.6
I could easily use CuedR every day	10	9.0	8.8	1.3
I could easily use CuedR every week	10	9.0	9.0	1.3

$F(17, 220) = 0.59$, $p = 0.87$. In addition, we conducted a post-hoc pairwise comparison using Tukey's HSD, which did not find any significant difference between any pair of portfolios in either registration time or in login time. These findings indicate that the usability in recognizing keywords did not vary significantly across different portfolios used in our study.

User Perception on the Efficacy of Different Cues

To understand user perception on the importance of different cues in aiding recognition, we asked them at the end of second session, "How often did the following cues assist you in recognizing keywords in CuedR?" In response, for each cue they selected one of five options: *Never*, *Rarely*, *Sometimes*, *Often*, or *Always*. Our results show that participants report using multiple cues to varying degrees to help recognize their keywords (see Figure 3). In particular, 92% of participants reported that the images were *always* or *often* helpful to recognize keywords, while 62%, 40%, and 14% of participants, respectively reported that spatial, phrase, and numerical cues were *always* or *often* helpful in recognizing keywords. The participants' diverse choices for cues to aid recognition and their high login success rate support our anticipation that letting users choose the appropriate cue(s) to their learning process aids the memorability for system assigned random passwords.

User Feedback on Usability and Applicability

We asked the participants to answer two sets of 10-point Likert-scale questions (1: *strong disagreement*, 10: *strong agreement*) at the end of the second session. We reversed some of the questions to avoid bias; the scores marked with (*) were reversed before calculating the modes, medians, and means. So, a higher score always indicates a more positive result for CuedR. To design the questionnaire, we carefully followed the guidelines provided in the existing password literature [9, 13, 44], including using nearly identical questions to those from other studies.

Usability

Participants showed a high degree of satisfaction with the usability (e.g., memorability, ease of login, ease of using either weekly or daily) of CuedR. Their feedback was also positive (mode, median, and mean higher than neutral) regarding login time, and they indicated that with practice they could log in quickly using CuedR (see Table 1). In our study, we could not test the usability of implicit feedback for CuedR, since most users did not make enough login mistakes to gain experience with it.

Applicability

At the end of second session, we asked 31 of the participants,[3] "Do you want to use CuedR in real life as a replacement to traditional textual passwords?" 84% responded 'Yes', 10% responded 'Maybe', and two participants responded 'No', where both of them mentioned that they would prefer traditional textual passwords in real life as they did not find any problems with them. User feedback about the applicability of CuedR in different online accounts is illustrated in Table 2.

Pilot study: Memorability for Multiple CuedR Passwords

It is common in password research to report a single-password study in the first article of a new authentication scheme, which helps to establish performance bounds and figure out whether multiple-passwords tests are worthwhile in future research. A recent survey [5] reported that out of 25 graphical password schemes proposed to date, only three have been evaluated through a multiple-password study, and none of these study results was reported in the first article. Since the use of multiple passwords is an important issue for deployment, however, we conducted a pilot study for multiple passwords, in addition to reporting the detailed results of a single-password study.

The study procedure was same as that in our single-password study, except that each participant was assigned three CuedR

Table 2. The applicability of CuedR for different online accounts. Scores are out of 10.

Online accounts	Mode	Median	Mean	SD
Bank	10	8.0	7.4	2.6
E-mail	10	9.0	8.1	2.1
Social Networking	10	9.0	7.7	2.4
University Portal	10	8.0	8.2	1.9
E-commerce	10	9.0	7.8	2.5

---

[3]We failed to ask the first six participants.

passwords (18 keywords, in total) instead of one. To administer this experiment, we created three different websites outfitted with CuedR and presented the sites to participants as tabs in an open browser window. Participants were free to select the order of websites at registration and login, but the tabs were arranged the same way every time. For this study, we recruited 11 students (9 men, 2 women) who came from various majors of our university. We believe that 11 represents a suitable sample size for a pilot study [22].

In this study, all of the participants were able to log in successfully within three attempts in both *login 1* (same day of registration) and in *login 2* (one week after registration). In *login 1*, nine participants (82%) succeeded on the first attempt for all three CuedR passwords. One participant (9%) succeeded to log in using two CuedR passwords on the first attempt, where she recognized 17 keywords on the first attempt and corrected the lone mistake on the second attempt. Another participant succeeded on the first attempt for one CuedR password, where she successfully recognized 15 keywords on the first attempt and corrected the mistakes on the second attempt.

In *login 2*, six participants (55%) succeeded on the first attempt to recognize all 18 keywords. Four participants (36%) successfully recognized 17 keywords on the first attempt, i.e., they succeeded to log in using two CuedR passwords on the first attempt. For another CuedR password, two (18%) of these four participants succeeded on the second attempt and other two participants succeeded on the third attempt. One participant (9%) successfully recognized 16 keywords on the first attempt. In particular, she succeeded to log in using one CuedR password on the first attempt and succeeded on the second attempt for other two CuedR passwords.

DISCUSSION

In this section, we discuss three important aspects of CuedR: i) impact, ii) acceptance, and iii) application. Here, the term *study* refers to our single-password study, unless otherwise specified. We conclude with a discussion on the scope for future research on CuedR.

CuedR: The Impact

Existing password systems fail to fully address users' cognitive limitations or leverage humans' cognitive strengths. Thus, despite a large body of research, it still remains a critical challenge to build an authentication scheme that provides both guessing resilience and high memorability. CuedR represents a breakthrough, offering high memorability for system-assigned random passwords, and shows a promising research direction to leverage humans' cognitive abilities for user authentication.

System-assigned passwords provide higher security against guessing attacks than user-chosen passwords, but it is difficult for most people to memorize them [18, 38, 49]. Users have varying cognitive strengths and abilities, and it is hard to know in advance what will help a given user to remember her password. In CuedR, we present a variety of visual, verbal, and spatial information related to randomly selected keywords in an organized way and then let users choose the

appropriate cue(s) to their learning process. Further, the cues can work together. When we asked users to identify the cues they used, 83.8% of users reported of using multiple cues often or always. As one participant commented, "The image, phrase and number naturally correspond in my mind, and make it easy to remember."

CuedR also shows that the cued-recognition class of password schemes, a new design point in the field, can be effective for user authentication. In particular, CuedR addresses each of the features needed in an effective graphical password scheme, as identified by Biddle et al. [5] from their comprehensive survey on the graphical password literature.

Here we mention a participant's feedback that particularly drew our attention: *"The multiple cues make it a helpful password scheme for autistic persons, who find it cognitively difficult to create secure passwords."* We appreciate such a thoughtful opinion and note it to be an important issue to be explored in future work.

CuedR: The Acceptance

In traditional user-chosen passwords, users bear the responsibility of ensuring security for their online account through a secure password that should be chosen with creativity and intelligence so that it achieves satisfactory memorability. For many users, this is a lot of work, and thus in many cases they compromise with security and create a weak but memorable password. A recent study [37] reveals that with the advancement of digital technology and widespread use of internet in recent years, users now better realize the importance of strong passwords than anytime before, and many of them intend to create secure passwords but just fail to achieve a good balance between security and memorability. So, rather than blaming users for predictable passwords, researchers should improve how authentication systems address human cognitive abilities.

Participants in our study seem to be convinced with the security provided by a system-assigned password. In a post-experiment open-ended question where they were asked about their opinion of CuedR, most of them reported high satisfaction with its security features.

Since the participants were convinced with the security of six system-assigned keywords, and could efficiently recognize each keyword in a reasonable time (6.3 seconds, on average) after a week of registration, they found the overall login time acceptable and reported satisfaction with the usability of CuedR (see Table 1). 84% of participants preferred to use the scheme in real life as a replacement to traditional textual passwords.

CuedR: The Applications

Although most of the participants reported strong agreement about using CuedR for all of the given account types (see Table 2), we must be cautious to recommend its application since textual passwords have lower login times than CuedR. Traditional textual passwords are fraught with security problems that make them less than desirable, especially for high-security accounts [14]. Ideally, there should be a clear sepa-

ration between the passwords used for low-security websites and high-security websites [7]. Thus, CuedR can be used as a standard authentication mechanism for online accounts with high security requirements and where logins occur relatively infrequently, such as financial (e.g., online banking, brokerage services) and e-commerce accounts [23]. A study by Hayashi and Hong [23] finds that users log into financial and e-commerce sites once a week on average, which is in agreement with the interval of one week before *login 2* in our study. By using CuedR for high-security accounts, it helps to build the mental separation with lower security accounts and avoid attacks based on password reuse and predictable patterns.

As compared to other cognometric graphical password schemes that present users with images only, the deployment of CuedR may require more effort, where separate portfolios of keywords are built accommodating both graphical and verbal cues. We note that each commercial deployment can use a small set of portfolios for all of its users. For example, with 10 portfolios, a phisher could correctly guess the first two portfolios for a user only 1% of the time.

Future Work
Now that lab-study results show promise for CuedR, it would be an interesting avenue for future work to evaluate the scheme through a long-term multiple-password study with larger and more diverse populations, where we would explore the training effect in reducing login time over more login sessions. A recent field study [2] reveals that login time decreases with the frequent use of a scheme due to training effects.

In the current interface of CuedR, users have to look at a separate table to find the phrase/fact related to a keyword (see Figure 1). In future work, we will test an alternate interface design to improve login time: The fact related to a keyword would be shown just below the graphical cue of that keyword, in which case users should require less time to find a phrase than finding it from a separate table.

While we have found that combining multiple cues shows promising results for authentication, we plan to pursue future studies to address the following issues: i) The impact of cues on the login performance of users from different age groups; ii) The usability of offering various combinations of cues; iii) The correlations between usability and the elimination of cue(s) from the interface over login sessions; iv) The usability of leveraging different cues from mobile devices.

CONCLUSION
In this paper, we present a novel authentication scheme, CuedR, which helps us to explore the efficacy of combining graphical, verbal, and spatial cues to improve the memorability of system-assigned random passwords. We also discuss the promise of CuedR in addressing the features of an effective graphical password scheme [5]. Although the login time is relatively high, our primary findings indicate high memorability for CuedR, suggesting that cued-recognition would be an important direction in password research to address the usability-security tension in authentication.

REFERENCES
1. Passfaces corporation. The science behind Passfaces. White paper, **http://www.passfaces.com/enterprise/resources/white_papers.htm.**

2. Al-Ameen, M. N., and Wright, M. A comprehensive study of the GeoPass user authentication scheme. Tech. rep., arXiv:1408.2852 [cs.HC], 2014.

3. Anderson, J. R., and Bower, G. H. Recognition and recall processes in free recall. *Psychological Review 79(2)* (1972).

4. Atinkson, C. R., and Shiffrin, M. R. Human memory: A proposed system and its control processes. *K.W. Spence and J.T. Spence (eds), Advances in the psychology of learning and motivation, New York academic press* (1968).

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

6. Bonneau, J. The science of guessing: Analyzing an anonymized corpus of 70 million passwords. In *IEEE S&P* (2012).

7. Bonneau, J., and Preibusch, S. The password thicket: Technical and market failures in human authentication on the Web. In *WEIS* (2010).

8. Bridis, T. Hacker impersonated Palin, stole e-mail password. Associated Press, Sep. 18 2008.

9. Chiasson, S., Biddle, R., and van Oorschot, P. C. A second look at the usability of click-based graphical passwords. In *SOUPS* (2007).

10. Chiasson, S., Forget, A., Biddle, R., and van Oorschot, P. User interface design affects security: Patterns in click-based graphical passwords. *International Journal of Information Security 8(6)* (2009).

11. Chiasson, S., Stobert, E., Biddle, R., and van Oorschot, P. Persuasive cued click-points: design, implementation, and evaluation of a knowledge- based authentication mechanism. *IEEE TDSC 9* (2012).

12. Chiasson, S., van Oorschot, P. C., and Biddle, R. A usability study and critique of two password managers. In *USENIX* (2006).

13. Chiasson, S., van Oorschot, P. C., and Biddle, R. Graphical password authentication using cued click points. In *ESORICS* (2007).

14. Das, A., Bonneau, J., Caesar, M., Borisov, N., and Wangz, X. The tangled web of password reuse. In *NDSS* (2014).

15. Davis, D., Monrose, F., and Reiter, M. On user choice in graphical password schemes. In *USENIX Security* (2004).

16. Dirik, A. E., Memon, N., and Birget, J.-C. Modeling user choice in the passpoints graphical password scheme. In *SOUPS* (2007).

17. Dunphy, P., and Yan, J. Do background images improve "Draw a Secret" graphical passwords? In *CCS* (2007).

18. Everitt, K., Bragin, T., Fogarty, J., and Kohno, T. A comprehensive study of frequency, interference, and training of multiple graphical passwords. In *CHI* (2009).

19. Florencio, D., and Herley, C. Where do security policies come from? In *SOUPS* (2010).

20. Forget, A. *A World with Many Authentication Schemes*. PhD thesis, Carleton University, 2012.

21. Forget, A., Chiasson, S., van Oorschot, P., and Biddle, R. Improving text passwords through persuasion. In *SOUPS* (2008).

22. Forget, A., Chiasson, S., van Oorschot, P. C., and Biddle, R. Persuasion for stronger passwords: Motivation and pilot study. In *PT* (2008).

23. Hayashi, E., and Hong, J. I. A diary study of password usage in daily life. In *CHI* (2011).

24. Hlywa, M., Biddle, R., and Patrick, A. S. Facing the facts about image type in recognition-based graphical passwords. In *ACSAC* (2011).

25. Jermyn, I., Mayer, A., Monrose, F., Reiter, M., and Rubin, A. The design and analysis of graphical passwords. In *USENIX Security* (1999).

26. Kaliski, B. RFC 2898: PKCS #5: Password-based cryptography specification version 2.0, Sep. 2000.

27. Li, Z., He, W., Akhawe, D., and Song, D. The emperor's new password manager: Security analysis of Web-based password managers. In *USENIX Security* (Aug. 2014).

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

29. Nelson, D. L., Reed, V. S., and McEvoy, C. L. Learning to order pictures and words: A model of sensory and semantic encoding. *Journal of Experimental Psychology: Human Learning and Memory 3(5)* (1977).

30. Nicholson, J., Coventry, L., and Briggs, P. Age-related performance issues for PIN and face-based authentication systems. In *CHI* (2013).

31. Paivio, A. *Mind and Its Evolution: A Dual Coding Theoretical Approach*. Lawrence Erlbaum: Mahwah, N.J., 2006.

32. Peters, M. Revised Vandenberg and Kuse mental rotations tests: Forms MRT-A to MRT-D. Tech. rep., Department of Psychology, University of Guelph, 1995.

33. Provos, N., and Mazieres, D. A future-adaptable password scheme. In *USENIX ATC* (1999).

34. Renaud, K. A visuo-biometric authentication mechanism for older users. In *British HCI* (2005).

35. Schaub, F., Walch, M., Konings, B., and Weber, M. Exploring the design space of graphical passwords on smartphones. In *SOUPS* (2013).

36. Schechter, S., Brush, A. J. B., and Egelman, S. It's no secret: Measuring the security and reliability of authentication via 'secret' questions. In *IEEE S&P* (2009).

37. Shay, R., Ion, I., Reeder, R. W., and Consolvo, S. "my religious aunt asked why i was trying to sell her Viagra": Experiences with account hijacking. In *CHI* (2014).

38. Shay, R., Kelley, P. G., Komanduri, S., Mazurek, M. L., Ur, B., Vidas, T., Bauer, L., Christin, N., and Cranor, L. F. Correct horse battery staple: Exploring the usability of system-assigned passphrases. In *SOUPS* (2012).

39. Shay, R., Komanduri, S., Durity, A. L., Huh, P. S., Mazurek, M. L., Segreti, S. M., Ur, B., Bauer, L., Christin, N., and Cranor, L. F. Can long passwords be secure and usable? In *CHI* (2014).

40. Shay, R., Komanduri, S., Kelley, P. G., Leon, P. G., Mazurek, M. L., Bauer, L., Christin, N., and Cranor, L. F. Encountering stronger password requirements: User attitudes and behaviors. In *SOUPS* (2010).

41. Silver, D., Jana, S., Boneh, D., Chen, E., and Jackson, C. Password managers: Attacks and defenses. In *USENIX Security* (Aug. 2014).

42. Stobert, E., and Biddle, R. Memory retrieval and graphical passwords. In *SOUPS* (2013).

43. Tari, F., Ozok, A., and Holden, S. A comparison of perceived and real shoulder-surfing risks between alphanumeric and graphical passwords. In *SOUPS* (2006).

44. Thorpe, J., MacRae, B., and Salehi-Abari, A. Usability and security evaluation of GeoPass: A geographic location-password scheme. In *SOUPS* (2013).

45. Tulving, E., and Thompson, D. M. Encoding specificity and retrieval processes in episodic memory. *Psychological Review 80(5)* (1973).

46. Tulving, E., and Watkins, M. Continuity between recall and recognition. *American Journal of Psych 86(4)* (1973).

47. Wickelgren, W. A., and Norman, D. A. Strength models and serial position in short-term recognition memory. *Journal of Mathematical Psychology 3* (1966).

48. Wiedenbeck, S., Waters, J., Birget, J., Brodskiy, A., and Memon, N. Authentication using graphical passwords: Effects of tolerance and image choice. In *SOUPS* (2005).

49. Wright, N., Patrick, A. S., and Biddle, R. Do you see your password? Applying recognition to textual passwords. In *SOUPS* (2012).

50. Yan, J., Blackwell, A., Anderson, R., and Grant, A. Password memorability and security: Empirical results. *IEEE Security and Privacy 2 (5)* (2004), 25.