Short Training Techniques to Enhance Usability of System-Assigned PINs

Israt Jahan Jui
Ontario Tech University
israt.jui@ontariotechu.net

Amirali Salehi-Abari
Ontario Tech University
abari@ontariotechu.ca

Julie Thorpe
Ontario Tech University
julie.thorpe@ontariotechu.ca

***Abstract*—Personal Identification Numbers (PINs) are widely used for authentication on mobile devices such as smartphones, which act as gateways to many important accounts (e.g., financial, email, etc.). Unfortunately, people tend to choose easy-to-recall PINs involving birthdays, anniversaries, or keypad patterns that are vulnerable to guessing attacks. System-assigned PINs can improve PIN security in this regard; however, they have usability problems such as feeling the need to store the assigned PIN, longer login times, and difficulty remembering. In this paper, we propose, design, and evaluate a set of short training techniques (16–34 seconds) inspired by implicit learning techniques, to improve the usability of system-assigned PINs. We evaluated our designs in a two-session user study with 184 university students. Our results show that some designs offer significant improvements in the login success rate, login times, and user perceptions. These advantages are in addition to our design's short single-session training, making it more compatible with typical registration workflows than previously proposed multi-session training techniques.**

***Index Terms*—Authentication, Personal Identification Numbers (PINs), Implicit Learning, Contextual Cueing (CC), Usability**

I. Introduction

PINs are used in a variety of computing devices such as mobile devices and tablets. These devices should be protected not only because they hold much of our personal and private information [1], but they are also being increasingly used as authentication gateways (via password managers, 2FA, or FIDO/passkeys [2]) to our online personal, financial, and professional worlds [3]. Many of these devices have email apps installed, which allow for many password and account resets. Many mobile devices and tablets use PINs as the primary authentication method. Even when biometrics are enabled, a PIN is often still enabled as a login option for when the biometric fails.

Unfortunately, user-chosen PINs often follow common patterns (e.g., birthdays or other memorable dates [4], [5]), leaving them susceptible to being guessed. Six-digit PINs have been proposed as a method to improve security, but unfortunately, they were found to offer little security advantages in practice, with some additional usability concerns (e.g., being slower to input [6]). The time taken to unlock mobile devices is already an annoyance to users [7], so this approach also has practical usability disadvantages. A system-assigned PIN, on the other hand, is randomly generated and ensures optimal security against guessing or credential-stuffing attacks. Banks used to assign such PINs to their customers to ensure heightened security. However, by the 1980s, they began allowing customers to select their PINs, primarily as a marketing strategy [8]. More recently, there have been recommendations suggesting a return to system-assigned PINs for enhanced safety [9].

System-assigned PINs often come with usability challenges that can undermine their security benefits. Users may feel compelled to write down these randomly generated PINs because they are difficult to remember, thus introducing new security risks. Additionally, longer login times [8] and the mental effort required to recall unfamiliar numbers can lead to frustration and decreased user satisfaction. In this work, we aim to address these usability challenges associated with system-assigned PINs.

Inspired by the implicit learning technique of *Contextual Cueing* [10], we have designed a set of novel short single-session training techniques for PIN user interfaces. Our designs aim to maintain the security of system-assigned PINs while improving its usability. Each of our designs were constructed to contain a subset of Contextual Cueing elements (including repetition, targets, and unique distractors), allowing us to assess the effectiveness of Contextual Cueing elements in improving usability. We conducted a user study ($N = 184$) to answer the following research questions: (1) Do our training designs improve the usability (e.g., login time, memorability, storage rate, or user perceptions) of system-assigned PINs? (2) As our designs are constructed to contain subsets of elements of the Contextual Cueing paradigm, which (if any) of these sets of elements appear to offer the most usability improvements?

Our contributions include: (i) The design of a set of novel training techniques for system-assigned PINs. These techniques are short (16-34 seconds) and take place in one session; (ii) The evaluation of our designs through a two-session study run with university students ($N = 184$); and (iii) An analysis of our designs, demonstrating considerable usability improvements for some. Our training time is also shorter than other training techniques that span several days [8]. We also discuss, based on our designs and their evaluation, which elements of the Contextual Cueing paradigm appear to hold promise in such training designs, and ways to further improve our specific training designs.

II. Related Work

We first review the general security and usability of both user-chosen and system-assigned PINs. We then discuss relevant training techniques for various types of authentication systems including system-assigned PINs to improve usability.

A. User-Chosen PINs

Many banking clients have been found to choose PINs based on birthdays or other memorable dates [4]. An analysis of 3.4 million four-digit PINs from leaked passwords found that "1234" was used by 10.7% of all PINs, followed by "1111" and "0000" [11], [12]. Moreover, keyboard patterns "2580", which is a "straight-shot" down the center of a keypad, and "across the corners" combinations are notable. Password datasets have also been found to contain many dates [13]. Users prioritize memorability over security when choosing PINs and sometimes repeat the same PIN for multiple assets [14]. This is problematic, as reusing PINs makes them vulnerable to credential-stuffing attacks. While security issues are present in user-chosen PINs across groups, there can be differences based on users' backgrounds (e.g., language or country of origin) [15]. Since user-chosen PINs are easy to remember, some people may think that upgrading four-digit PINs to six-digit PINs can solve the security issue. However, 6-digit PINs provide minimal security advantages which are not worth the usability losses such as being slower to input and harder to remember [6].

B. System-Assigned PINs

Though system-assigned PINs are more secure, they suffer from usability problems such as users feeling the need to write them down and requiring longer login times [8]. Older adult users face the most difficulty remembering their system-assigned PINs [16]. Research on number-chunking techniques, which split longer PINs into smaller groups (e.g., 480271 as 48-0271), examined the memorability of system-assigned PINs. Success rates were 74% for 4-digit and 55% for 6-digit PINs after two days, with average login times of 22.6 and 35.5 seconds, respectively [9].

C. Training Techniques for Authentication

Other studies have examined various training techniques to improve the usability of system-assigned authentication secrets. Not all techniques are equally successful; some research [17] found that about half of people prefer to stick to their own memorization strategy. The existing training techniques fall into two categories.

Repetition-based Training. Spaced repetition training techniques have been used to improve the memorability of system-assigned PINs [8]. Two designs were explored: *Second PIN*, which adds a secondary numeric PIN after the user's self-chosen PIN during login, and *Mapping*, which changes the numeric keypad layout at each login. Training involved 25 logins over 2-8 days, with median learning times of 81 seconds for Second PIN and 172 seconds for Mapping in one experiment, and a median of 40 seconds for Second PIN and 117 seconds for Mapping in a second experiment. While most participants successfully memorized their PINs, 10% failed to do so with the Mapping design. Spaced repetition has also been studied for passwords [18] and passphrases [19]. While the spaced repetition method is promising for memorability, its requirement of a longer training session (i.e., about 1-3 minutes for PINs) that is done over multiple sessions is a practical disadvantage to its usability and adoption.

Implicit learning-based Training. Implicit learning is the non-episodic learning of complicated knowledge unconsciously, with no awareness of what has been acquired [20]. Some graphical authentication techniques explored priming-based implicit learning at registration time, which involved testing the accuracy of responses to a set of challenge images [21], [22]. Another technique aims to employ implicit learning techniques in order to resist coercion attacks [23]. The idea is interesting as users cannot be compelled to disclose the secret since they don't know it consciously, however, it has long (30-45 minute) training times and only 47% of participants could successfully authenticate one week later. Contextual Cueing (CC) and Semantic Priming (SP) techniques were used to improve recall of system-assigned passphrases [24]. Combining these methods (CC-SP) led to higher recall rates, with 88% after one week (compared to 57% in the control group). However, its security was similar to system-assigned PINs but it offers limited practicality as a PIN replacement due to its requirement to be able to modify the screen layout.

Our work aims to address usability challenges of system-assigned PINs, by combining repetition and implicit learning-based training approaches. Our UI design prioritizes reducing training time to under a minute in a single session (for compatibility with registration time), contrasting with prior methods requiring one to several minutes of interaction over multiple days. Additionally, we introduce a novel repetition-based training UI and are the first to explore implicit learning through Contextual Cueing combined with repetition approaches for training system-assigned PINs.

III. Background

As it is relevant to our designs, we review implicit learning, focusing on Contextual Cueing (CC). Implicit learning occurs unconsciously, often through repeated exposure or experience. For example, balance on a bicycle is learnt implicitly, as it is done by trial and error. In contrast, explicit learning occurs through conscious and deliberate effort, such as in learning how to use the handbrake on a bicycle for the first time.

Contextual Cueing is a phenomenon where people become faster at finding a target in a visual search because the arrangement of objects in the scene is occasionally repeated [10]. The term *context* refers to the 2D arrangement of objects in visual displays [10]. For example, if a target always appears in the same spot within a familiar arrangement of objects, people locate it more quickly compared to new, unfamiliar arrangements—even if they don't consciously notice that the setup is repeated. This implicit learning of the target's location

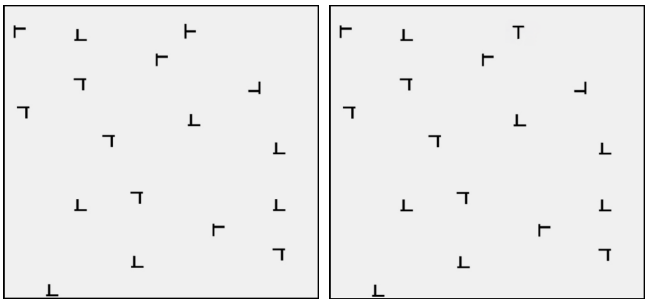


Fig. 1: Example repeated contexts used in classical Contextual Cueing, where the target is the 'T' (which may change orientation between repetitions, as shown here), and the shapes like an 'L' are the distractors.

within the scene occurs through repeated exposure to the specific spatial arrangement on the scene (aka *context*), which guides attention, making it easier to locate the target. This in turn reduces search time or reaction time to locate the target. Contextual Cueing is robust across various populations, including young children, older adults, and individuals with autism spectrum disorders [25].

In a Contextual Cueing experiment, participants search for a target (e.g., a letter 'T') in a 2D arrangement of distractors (e.g., the letter 'L'). The 2D arrangement is called a *context*, an example of which is shown in Figure 1. If the spatial arrangement of distractors is repeated, participants locate the target faster without realizing the context aided their search. Typically, the experiment will show a series of novel (previously unseen) contexts, and repeated (previously seen) contexts. After some number of repetitions, the reaction time of repeated contexts will become faster.

We hypothesize that Contextual Cueing can be used to train users on system-assigned PINs, as it has been with passphrases [24].

IV. PIN Training System Designs

Our training system designs are, to the best of our knowledge, the first to attempt incorporating implicit learning techniques in training system-assigned PINs. We describe our design of different systems to evoke Contextual Cueing (see Section IV-A). We implemented another training system that only uses simple repetition (see Section IV-B). The goal of studying these designs is to determine which elements of Contextual Cueing are most effective. As summarized in Table I, our training designs incorporate different combinations of three key components: repetition, targets, and unique distractor arrangements. The CC design employs all three elements—repetition, targets, and unique distractor arrangements. The CC-RP design includes repetition and targets but does not use unique distractor arrangements. The RP design relies solely on repetition.

Design Choices. Our goal in each design is to create a single short training session to improve usability. Each of these designs asks the user to input their system-assigned PIN 5 times. The decision to require 5 repetitions is based on other work on Contextual Cueing [24] and our early verification in pilot testing that the input time continued to reduce until about 4 repetitions; the plateau of input time is considered an indication that learning has taken place. Users have 30 seconds to input each digit, or the training will time out and require restarting. When users incorrectly input a digit, the system will tell them it is incorrect.

TABLE I: Different elements in each training design. The designs are used to compare the impact of Contextual Cueing elements (i.e., repetition, targets, and unique distractors.)

Design	Repetition	Targets	Unique Distractors
CC	✓	✓	✓
CC-RP	✓	✓	
RP	✓		

For all of the following designs, with the exception of CC, the login UI is the same as one would typically see on a PIN login screen, as shown in Figure 5(b).

A. Contextual Cueing (CC) Designs

There are some inherent constraints in mapping Contextual Cueing to a PIN system: (a) We cannot replace the PIN digits with the targets and distractors of classical Contextual Cueing experiments—the keys must contain numbers. (b) We cannot change the position of each key to create different 2D arrangements, as the PIN pad has limited screen space on mobile devices. (c) While we could alter the position of each digit on the PIN pad, we believe this would lead to confusion as users have already learned where to expect each digit's positions in the conventional layout of these numbers. Therefore, our designs maintain our goal of keeping the arrangement of digits using the traditional PIN pad layout.

In the CC training phase, a user must search to locate the target (in our case the system-assigned digit) among a set of distractors (in our case the other digits). Thus, the target should be not easy to find, but possible for a user who is paying attention. Our design tilts the target digit (20 degrees) and asks users to select it. Each of the 4 system-assigned digits is presented in this way, in sequence, 5 times. The goal is that by the end of the 5 training repetitions of each digit, the user will learn the position of each of their 4 target digits in relation to the other elements of the PIN pad.

We show how we mapped Contextual Cueing targets and distractors to a PIN display in Figure 2. For a given display, the target, or item the user is searching for, is the PIN digit they are assigned. The assigned target PIN digit is signaled to the user by changing its orientation (see Figure 2)—this was decided after iterative pilot testing of different fonts and different orientations. For a given display, the distractors, or items that the arrangement of distinguish the display from the others, are an overlay that shades 4 out of the 10 PIN digits on the screen (see Figure 2). The goal is to keep the distractors as simple as possible. Each display has a different target (tilted

digit the user needs to find) and different set of distractors (orange shaded PIN digits).

- **CC** presents a sequence of 4 screens, where for each users are asked to tap on the number with different orientation. Each screen shows one digit of the system-assigned PIN on a regular PIN pad (see Figure 3(a)). The sequence of 4 screens is shown 5 times. CC also aims to create different distractor arrangements for each of the 4 system-assigned digits.

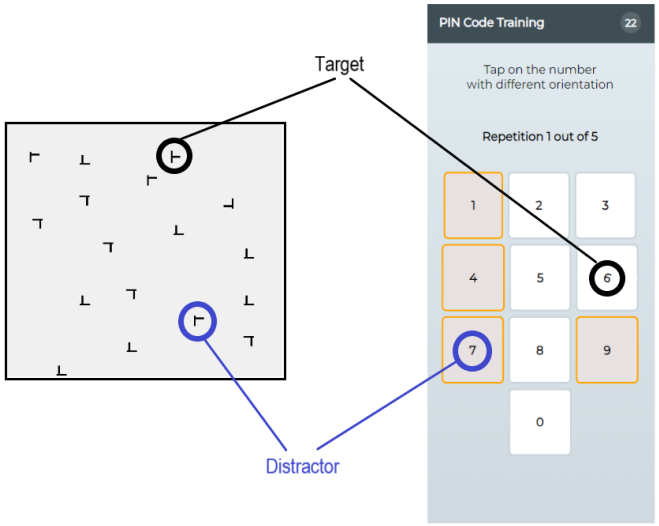
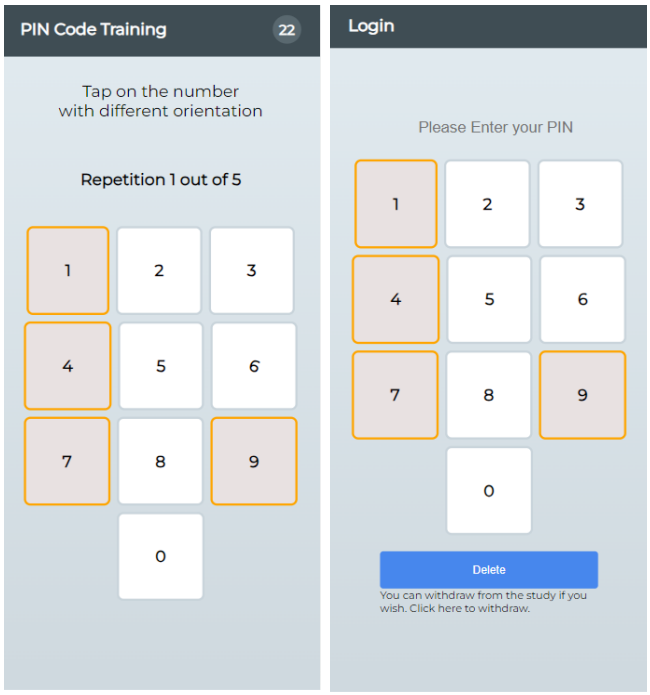


Fig. 2: Mapping between classical Contextual Cueing and our CC design for PIN training, indicating which elements are the target and distractors.

- **CC-RP**, like CC, presents a sequence of 4 screens, where for each users are asked to tap on the number with different orientation. Each screen shows one digit of the system-assigned PIN on a regular PIN pad (see Figure 4(a)). The sequence of 4 screens is shown 5 times.
  However, CC-RP differs from CC as it does not have the orange overlay; instead, each screen for the four PIN digits is the same standard PIN display. Our reason for this design was to help evaluate whether the overlay of distractors used in our CC design was useful or not. Additionally, the user's system-assigned PIN is also shown at the top of the screen (see Figure 4(b)). The objective of this strategy was to reinforce learning of the PINs by utilizing both the implicitly learned target locations and the repetitive exposure to viewing the entire PIN.

B. Repetition (RP)

In this design, users were simply asked to enter the system-assigned PIN five times (see Figure 4(b)). In this condition, there was no implicit learning technique involved. The purpose of this is to determine whether simple repetition in a single short session, without any change to the UI, could offer an improvement.



(a) CC Training (b) CC Login

Fig. 3: CC's (a) training and (b) login screens for the same digit. In the training phase, the user must find and select the digit with different orientation (in this example '6'). Note that each of the 4 digits in the system-assigned PIN has a different overlay pattern of orange digits. Observe how the orange overlay patterns remain for login, but the tilting of digits used for training is not present during login.

V. Methodology

We aim to evaluate whether our training designs (see Section IV) for learning and communicating system-assigned PINs improve usability through a study run with students in our university. In addition to our target designs of CC, CC-RP, and RP, our study also had a control group that did not involve any training or reinforcement; they were asked to type the given PIN once (see the Control group UI in Figure 5). For our study, we implemented our designs on a website that only allowed access from mobile devices.

A. Study Tasks and Structure

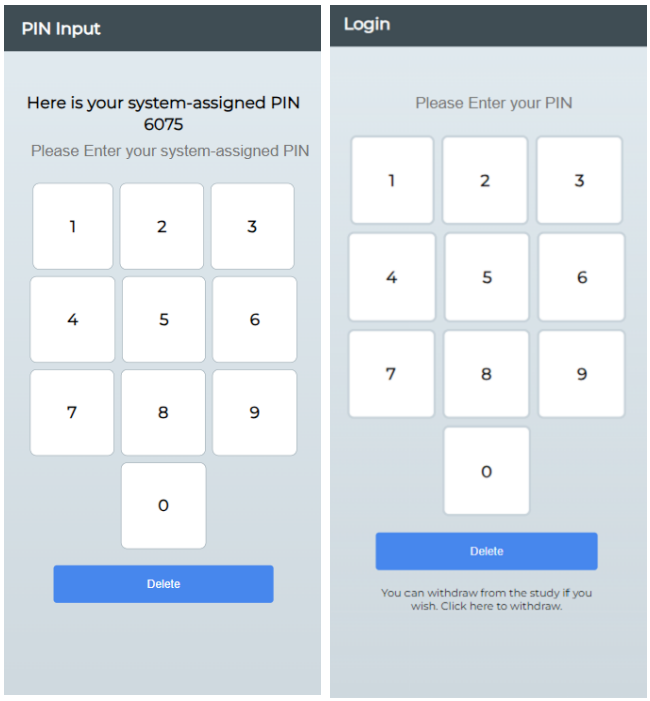
Our study was approved by our university's research ethics board and consisted of two sessions: a registration session (involving our training designs, where applicable), and a login session 24-48 hours later. We detail the task structure of each session below.

Session 1 (registration). This session aims to provide a registration phase using our training designs, asking users to login, as well as collecting user perceptions. This session includes these tasks in the order shown below:

1. *Consent form.* Participants were asked to read and sign a consent form; after they agreed, they were randomly

(a) CC-RP (b) RP

Fig. 4: Training UIs for (a) CC-RP and (b) RP. The user must input their correct PIN 5 times to complete training. Notice the difference is that CC-RP has one PIN digit with different orientation per screen (in this example, the number '5'). The login UI for both of these designs is the same as a typical PIN login screen (see Figure 5).



(a) Registration (b) Login

Fig. 5: Control group UIs.

assigned to one of 4 groups (three training designs, and one control). They were informed there will be a second session in 1-2 days.
2. *PIN assignment.* Users were assigned a randomly generated four-digit PIN. Users outside the control group received a brief training session with one of our designs, as described in Section IV.
3. *Demographic questionnaire.* We ask 5 demographic questions (age, gender, education, primary area of study or work, and first language).
4. *Login.* Users were asked to log in with the assigned PIN. If users were unable to remember the assigned PIN, they were given the option to redo the training (applicable for CC, CC-RP, and RP) or see the PIN again (Control group only). Note this rarely occurred as described in Section VI.
5. *Feedback questionnaire.* We ask for feedback on the system and its training session.

Session 2 (24-48 hours after Session 1): This session aims to test memorability and login usability of the system-assigned PIN, and collect user perceptions, under the different training conditions. It was held after 1-2 days of Session 1 and includes these tasks in the order shown below:

1. *Login.* Users had 5 attempts to input their correct PIN. If a user fails to input the correct PIN within 5 attempts, they are considered to have forgotten the PIN.
2. *Final questionnaire.* This includes questions related to their recollection of training, assigned PIN, and the system usability.

B. Recruitment and Compensation

Students from our university were recruited via broadcast email to all students. They were randomly assigned to one of 4 groups: CC, CC-RP, RP, or Control. To encourage a high participation rate, participants who completed Session 1 were entered into a draw to win 1 of 2 $100 bank deposits, coffee gift cards, or Amazon gift cards of their choice. Upon completion of Session 2, they were provided their choice of a $5 bank deposit, coffee gift card, or Amazon gift card.

C. Participant Confidentiality

The only personal data we collected was the email address of the students in order to send them rewards via their emails. Once we sent them the rewards, we removed their email address from our database and only use an anonymized ID to distinguish participants. No mappings between anonymous IDs and email addresses were retained.

VI. Results

Here we describe our participant demographics and dropouts, compare various usability metrics for our training designs, and analyze factors that may help inform future training designs.

A. Participant Demographics

We recruited *N=201* students from our university, *184* of whom completed the study. To reduce impacts of social desirability bias and any perceived pressure, students in the program the researchers teach in were excluded from participation. Across all groups, more participants identified as female (61–66%) than male (30–36.5%). The majority of participants were between the ages of 18 and 25. Most participants had a high school degree (60–68%), followed by a bachelor's degree (22–29%) as their highest education to date. Field of study was divided among many disciplines. English was the first language for most (65.3%–72%).

Participation verification and dropout. To prevent fraudulent participation, we required valid university student email addresses for compensation. After 24 hours of completing Session 1, participants were invited via email to join Session 2. Of the 201 participants recruited, 184 returned for Session 2, resulting in an 8.46% dropout rate.

B. Training Times

Training time significantly impacts the usability and adoption of any training approach for system-assigned PINs. Long or cumbersome training sessions, even if they are one-time, hinder user adoption and perception. Figure 6 shows the training times of each of our designs. The median training times are 34s for CC, 20s for CC-RP, and 16s for RP. As expected, this is substantially less than the time for spaced repetition (which has median times between 40 seconds and 3 minutes [8]). Our training designs can be done in a single session at registration time, whereas spaced repetition requires repetitions over multiple sessions; as such, our designs are more compatible with typical account registration processes.

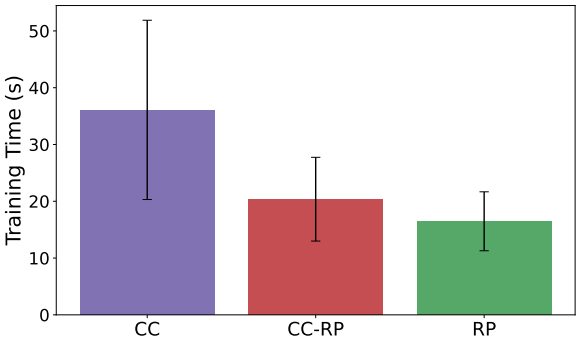


Fig. 6: Average training times for each training system.

To determine whether training times might be predictive of whether or not users will remember their PINs, we also analyzed the training times for users who remembered the PINs vs. who did not. We found that users who forgot their PINs, took more time to take the training than the users who remembered the PINs. This might point towards a way to improve training by providing more repetitions if the training time for each repetition hasn't yet decreased as expected.

C. PIN Storage Rates

In the Session 2 questionnaire (at the end of the study), we emphasized that participants could choose to record their allocated PIN and encouraged them to inform us if they did. Our system also detected instances of PIN copying. Table II indicates the storage rates per group. While the results show all three training methods reduced PIN storage rates compared to the Control group, the improvement was not quite enough to reach statistical significance ($\chi^2 = 5.47$, $p = 0.14$, $df = 3$).

For the remainder of our analysis, we excluded data from participants who reported writing down their PIN or were detected as having copied it.

TABLE II: PIN storage rates. PINs are considered stored if copy/paste was detected or participants specified they recorded their PIN.

Group	Stored PINs
Control	13/43 (30%)
CC-RP	8/47 (18%)
RP	9/49 (17%)
CC	5/45 (11%)

D. Memorability

We measure the memorability explicitly by two different measures (1) the recall rate, capturing the percentage of users who could remember their assigned PINs within five login attempts, and (2) the login success rate, capturing the number of login attempts required for those users who could eventually remember their assigned PINs within five login attempts.

1) Recall Rate: The Session 1 recall test was performed immediately after the short demographic questionnaire (about one minute after training). Although we do not consider this our main recall test, for completeness we report on its data here: only one participant using CC forgot the PIN and took the training one additional time. One participant in the Control group forgot their PIN once as well. No participants using RP or CC-RP forgot their PIN in Session 1.

The Session 2 recall test was performed 24-48 hours after being assigned the PIN in Session 1. The recall rates within 5 attempts for each system are comparable: CC has the highest recall rate of 85%, followed by RP with 82.50%, then CC-RP and Control with 80%.

2) Login Success Rate : Login success rate is an indication of memorability as it captures the ease of recalling the PIN. Login success rates are reported in Table III as the number of total successful login attempts divided by the number of total login attempts for those participants who eventually recalled their PINs. The difference in login success rates was significant ($\chi^2 = 10.35$, $p = 0.016$, $df = 3$) among the four groups. Post-hoc examination of residuals indicated this difference is mostly due to a significantly worse performance in Control vs. a significantly better performance in CC-RP, suggesting that CC-RP has a positive impact on memorability. In Figure 7 we visualize the number of attempts required for successful

logins. CC-RP had less than 10% of participants with failures, each of whom only had one failure each, whereas Control had nearly 40% of participants with failures, many of whom had more than one. This further indicates that CC-RP reduces the number of login errors, particularly relating to the first login attempt.

TABLE III: Session 2 percentage of login attempts that were successful, for those who eventually recalled their PIN.

Group	Login Success Rate
CC-RP	94%
RP	77%
CC	74%
Control	62%

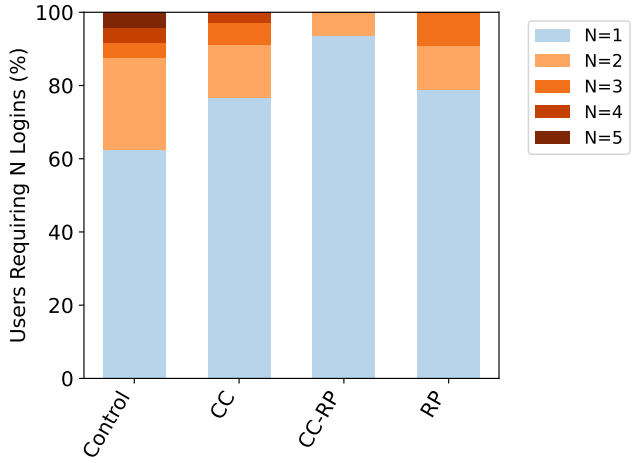


Fig. 7: Distribution of participants requiring N login attempts.

E. Login Time

We measure login time from when the login page loads until the time the user successfully logs in, including any time spent on login failures. Since login failures are included in this metric, this metric also indirectly captures memorability. Login time is an important usability metric that relates to ease of use and user satisfaction. Users could become impatient if it takes too long, which can lead to users choosing their own, less secure PIN.

Groups	Session 1	Session 2
Control	3.87 ± 3.32	26.94 ± 56.56
CC	4.76 ± 2.88	8.19 ± 06.43
CC-RP	3.45 ± 2.24	7.52 ± 10.54
RP	3.48 ± 1.64	6.77 ± 09.74

TABLE IV: Average login times (± standard deviation) prior to successful login. The times include time spent on login failure.

Table IV shows the average login times for both sessions. For Session 1, login times are consistent and comparable. However, for Session 2, the login time for all training-based groups of CC, CC-RP, and RP was lower than Control (see Table IV). To determine if our training designs improve login time, we performed a one-way ANOVA, which found a statistically significant difference in mean login time between at least two groups ($F(3, 118) = 3.397$, $p<0.05$, η^2=0.08). A Tukey HSD post hoc analysis indicated that the Control group differed significantly vs. all groups using our training designs and that no other groups differed significantly from each other. Figure 8 shows how CC, CC-RP and RP have significantly lower login times compared to Control.

These findings suggest that our training designs have a positive impact by lowering login times.

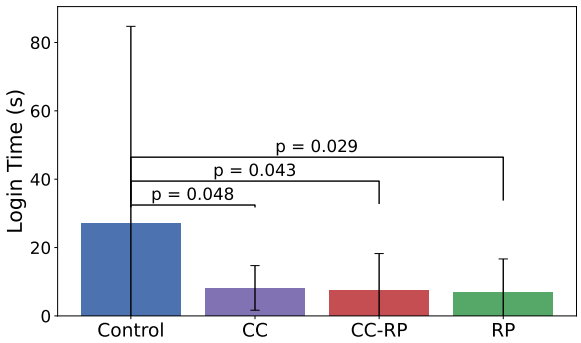


Fig. 8: Average login times for Session 2. The p-values shown are those that are significant (< 0.05) according to Tukey's HSD.

F. Usability Perceptions

At the end of Session 2, we asked participants to rate their agreement with the statement 'I thought the system was easy to use.' Figure 9 shows that participants strongly agreed that the RP and CC-RP systems were easy to use. There was more agreement that these systems were easy to use than for Control. The CC system was rated least easy to use.

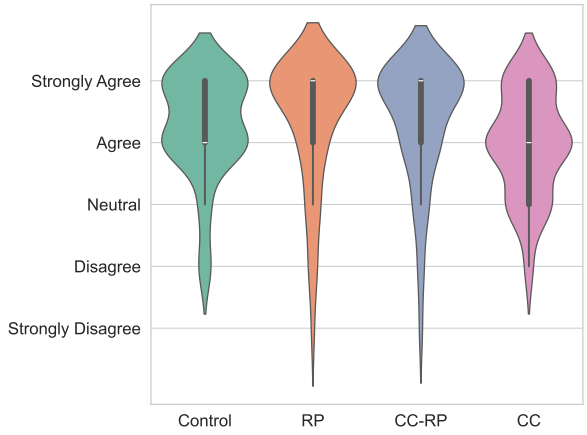


Fig. 9: Perceived usability as measured by agreement with the statement 'I thought the system was easy to use' in Session 2.

VII. Discussion

We discuss our student study's findings and how they relate to our research questions herein.

A. RQ1: Do Our Designs Improve Usability?

Our results suggest that our CC-RP design improves usability in terms of perceived usability, reduced login times, and higher login success rates, with low training time. RP had positive results in terms of perceived usability, reduced login times, and low training time, but it did not have the same improvement as CC-RP in terms of login success rates. Surprisingly, we found our CC design only had positive results for reducing login times—it does not appear to offer improvements in perceived usability, or login success rates. CC also had the longest training time of the three training designs.

1) Login Times: We found that all training designs had significantly lower login times than the control group. Since the login times include incorrect login attempts, this metric does somehow capture aspects of memorability too. Login time is especially critical for PIN-based systems, as PINs on mobile devices are used frequently, for example, users unlock their phones about 50 times per day on average [7]. Faster login times not only enhance user satisfaction and system efficiency but also reduce frustration, making the system more practical for everyday use. Our results suggest that the training designs contribute to a smoother and more user-friendly login experience.

2) Memorability: While all groups had comparably high recall rates, CC-RP had the lowest number of login errors, as only two participants had any login failures at all, and even they only had one each. This is shown in Figure 7 and further measured by the significantly higher login success rate (94%) vs. the Control group's 62% login success rate. Overall, it appears that the main advantage of CC-RP is that it reduces the number of login failures among those who recall the PIN. This finding is further supported by its reduced login times. We found all training designs reduced storage rates vs. Control; however, the effect was not significant after correction. Further study with larger sample sizes may help determine whether these differences are significant.

3) Usability Perceptions: CC-RP and RP both appear to improve perceived ease of use; however, CC does not.

B. RQ2: Which Contextual Cueing Elements Show Promise?

As discussed in Section IV-A, our training designs incorporate different combinations of three key elements of Contextual Cueing: repetition, targets, and unique distractor arrangements. The CC design employs all three elements (repetition, targets, and unique distractor arrangements). The CC-RP design includes repetition and targets but does not use unique distractor arrangements. The RP design relies solely on repetition.

Our findings indicated that RP appeared to offer some usability benefits alone, but did not improve login success rate like CC-RP. This finding indicates that adding targets to the training is useful. However, the addition of unique distractor arrangements, as we implemented in CC, did not offer the same benefits to login success rates, nor did it improve perceived ease of use. Another indicator that CC did not perform well is that the training time took much longer than for the CC-RP and RP designs. The free-form comments offer some insights into reasons for this, where a few participants indicated that they mistakenly assumed the distractors were the assigned digits. Taken together, these findings support that the CC design's implementation of unique distractors detracted from the benefits gained in CC-RP design (i.e., the repetition and fixed targets). We recommend that any future attempts at designing unique distractors ensure that the distractors do not focus on the digits themselves, to avoid potential confusion about which digits are part of the assigned PIN.

C. Potential Training Improvements

Beyond the total training time, we analyzed the training data to determine whether learning had taken place as intended.

We found that after the first repetition, the training time dropped and continued to decrease until it began to plateau between repetitions 3 and 4. This is consistent with research on implicit learning training for passphrases [24]. Interestingly, we noticed that participants who took a longer time to learn often couldn't remember their PINs in the next session. This might point towards a way to improve training by providing more repetitions if the time hasn't yet decreased as expected.

D. Interpretation

Our results suggest that our CC-RP design takes a valuable step towards improving the usability of system-assigned PINs in many ways (login time, login success, and user perceptions). RP also offers benefits for login time and user perceptions, but has more subtle improvements for memorability. CC also offers login time improvements, but suffers in terms of user perception and longer training times. It appears that the targets added to the training UI for CC-RP were useful, but the distractors added for CC were perhaps too distracting.

The median training times for RP and CC-RP were best (16 and 19 seconds respectively). These are less than half the time of the spaced repetition training approach [8], the fastest design for which takes a median time of 40 seconds over multiple sessions. Compared to spaced repetition, our training techniques are much quicker and it takes only a single training session with 5 repetitions.

Our results point towards possibilities for even further improvements. We found that most of the CC-RP recall errors were due to incorrect order despite remembering the correct 4 digits. This was a common error in all of the conditions, so there may be an opportunity to design training systems that improve on incorrect PIN orders.

E. Limitations

Our study was run online with university students. While studying this group had advantages of having assurance each participant's account is linked to an attentive human, this

population can be concerned about their performance in the eyes of their teachers or teaching assistants (i.e., a social desirability bias). Although this was mitigated to some degree as students in the program the researchers teach in were not permitted to participate in our study, university students remain a population who are typically more tech-savvy and fast. This means their training and login times may be faster than other populations. In particular, this may have boosted the performance of our Control group in various metrics. For example, it may be the reason for a lower storage rate of our Control group (30%) compared to the 45% of other studies that used crowdsourcing [8]. Also, the 80% recall rate for the Control group is higher than one might expect for system-assigned PINs. Our results should be interpreted in light of our sample population characteristics.

VIII. Concluding Remarks and Future Work

Our work takes a practical step towards improving the usability of system-assigned PINs. Our three designs involve only one short (16-34 second) training session, which is compatible with typical account registration processes. Two of our designs (CC-RP and RP) significantly improve both login times and user perception. Our CC-RP design additionally improves memorability in terms of login success rate. These results lead us to conclude that the repetition and target elements of our training designs were both useful. However, the distractors used in our CC design appear to be too distracting. We suggest that future designs should employ the techniques of CC-RP, but experiment with different approaches to distractors than our implementation of CC. For example, it might be possible to use a background image or pattern behind the digits, so that the interface looks different but doesn't imply that people need to interact with the distractors themselves.

Future work also includes personalizing the training, as some people may benefit from more repetitions. To determine whether training times might be useful in predicting users who will not recall their PIN, we have also analyzed the training times for users who remembered the PINs vs. who did not. To summarize, we found that users who forgot the PINs took more time to complete the training than the users who remembered their PINs. This might point towards a way to improve training by providing more repetitions until the repetition time decreases as expected.

Future research should aim for larger and more diverse participant samples (e.g., involving older adults and younger children). It may also experiment with using these training techniques for longer system-assigned PINs. Another relevant future direction is to explore influencing users to choose PINs that are more random.

Acknowledgments

This research was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC).

References

[1] B. Martínez-Pérez, I. De La Torre-Díez, and M. López-Coronado, "Mobile health applications for the most prevalent conditions by the world health organization: review and analysis," *Journal of Medical Internet Research*, vol. 15, no. 6, p. e120, 2013.

[2] S. G. Lyastani, M. Schilling, M. Neumayr, M. Backes, and S. Bugiel, "Is FIDO2 the kingslayer of user authentication? a comparative usability study of FIDO2 passwordless authentication," in *2020 IEEE Symposium on Security and Privacy (SP)*. IEEE, 2020, pp. 268–285.

[3] P. R. Center, "Demographics of mobile device ownership and adoption in the united states — pew research center," https://www.pewresearch.org/internet/fact-sheet/mobile/, April 2021, (Accessed on 08/11/2023).

[4] J. Bonneau, S. Preibusch, and R. Anderson, "A birthday present every eleven wallets? The security of customer-chosen banking PINs," in *Proceedings of the International Conference on Financial Cryptography and Data Security*, 2012.

[5] S. Davidson, "RBC refuses to refund ontario woman $8,772 because of PIN — CTV news," https://toronto.ctvnews.ca/ontario-woman-warns-about-choosing-credit-card-pin-after-rbc-refuses-to-refund-8-772-1.5895738, 2022, (Accessed on 01/14/2023).

[6] C. W. Munyendo, P. Markert, A. Nisenoff, M. Grant, E. Korkes, B. Ur, and A. J. Aviv, ""The same PIN, just longer": On the (in)security of upgrading PINs from 4 to 6 digits," in *In Proceeding of the 31st USENIX Security Symposium*, 2022.

[7] M. Harbach, E. Von Zezschwitz, A. Fichtner, A. De Luca, and M. Smith, "It's a hard lock life: a field study of smartphone (un)locking behavior and risk perception," in *Proceedings of the 10th Symposium on Usable Privacy and Security (SOUPS)*, 2014.

[8] S. Schechter and J. Bonneau, "Learning assigned secrets for unlocking mobile devices," in *Proceedings of the 11th Symposium On Usable Privacy and Security (SOUPS)*, 2015.

[9] J. H. Huh, H. Kim, R. B. Bobba, M. N. Bashir, and K. Beznosov, "On the memorability of system-generated PINs: Can chunking help?" in *Proceedings of the 11th Symposium On Usable Privacy and Security (SOUPS 2015)*, 2015.

[10] M. M. Chun and Y. Jiang, "Contextual cueing: Implicit learning and memory of visual context guides spatial attention," *Cognitive psychology*, vol. 36, no. 1, pp. 28–71, 1998.

[11] DataGenetics, "PIN number analysis," https://www.datagenetics.com/blog/september32012/index.html, 2012, accessed: 01/13/2023.

[12] T. Guardian, "The most common PIN numbers: Is your bank account vulnerable? — debit cards," https://www.theguardian.com/money/blog/2012/sep/28/debit-cards-currentaccounts, 2012, accessed: 01/13/2023.

[13] R. Veras, J. Thorpe, and C. Collins, "Visualizing semantics in passwords: The role of dates," in *Proceedings of the 9th International Symposium on Visualization for Cyber Security (VizSec)*, 2012.

[14] H. Khan, J. Ceci, J. Stegman, A. J. Aviv, R. Dara, and R. Kuber, "Widely reused and shared, infrequently updated, and sometimes inherited: A holistic view of PIN authentication in digital lives and beyond," in *Annual Computer Security Applications Conference (ACSAC)*, 2020.

[15] D. Wang, Q. Gu, X. Huang, and P. Wang, "Understanding human-chosen PINs: Characteristics, distribution and security," in *Proceedings of the 2017 ACM on Asia Conference on Computer and Communications Security*, 2017.

[16] J. Nicholson, L. Coventry, and P. Briggs, "Age-related performance issues for PIN and face-based authentication systems," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2013.

[17] K. Renaud and M. Volkamer, "Exploring mental models underlying PIN management strategies," in *Proceedings of the 2015 World Congress on Internet Security (WorldCIS)*, 2015.

[18] J. Blocki, S. Komanduri, L. Cranor, and A. Datta, "Spaced repetition and mnemonics enable recall of multiple strong passwords," in *Network and Distributed System Security Symposium (NDSS)*, 2015.

[19] J. Bonneau and S. Schechter, "Towards reliable storage of 56-bit secrets in human memory," in *Proceedings of the 23rd USENIX Security Symposium*, 2014.

[20] C. A. Seger, "Implicit learning." *Psychological bulletin*, vol. 115, no. 2, p. 163, 1994.

[21] C. Castelluccia, M. Dürmuth, M. Golla, and F. Deniz, "Towards implicit visual memory-based authentication," in *Proceedings of the Network and Distributed System Security Symposium (NDSS)*, 2017.

[22] T. Denning, K. Bowers, M. Van Dijk, and A. Juels, "Exploring implicit memory for painless password recovery," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2011.

[23] H. Bojinov, D. Sanchez, P. Reber, D. Boneh, and P. Lincoln, "Neuroscience meets cryptography: Crypto primitives secure against rubber hose attacks," *Communications of the ACM*, 2014.

[24] Z. Joudaki, J. Thorpe, and M. V. Martin, "Reinforcing system-assigned passphrases through implicit learning," in *Proceedings of the ACM Conference on Computer and Communications Security (ACM CCS)*, 2018.

[25] Y. V. Jiang and C. A. Sisk, "Contextual cueing," in *Spatial Learning and Attention Guidance*, S. Pollmann, Ed. New York, NY: Springer US, 2020, pp. 59–72. [Online]. Available: https://doi.org/10.1007/7657_2019_19