PiXi: Password Inspiration by Exploring Information

Shengqian Wang(✉), Amirali Salehi-Abari, and Julie Thorpe

Ontario Tech University, Canada
shengqian.wang@ontariotechu.net, {abari,Julie.Thorpe}@ontariotechu.ca

Abstract. Passwords, a first line of defense against unauthorized access, must be secure and memorable. However, people often struggle to create secure passwords they can recall. To address this problem, we design ***P**assword **i**nspiration by e**X**ploring **i**nformation (PiXi)*, a novel approach to nudge users towards creating secure passwords. PiXi is the first of its kind that employs a password creation nudge to support users in the task of generating a unique secure password themselves. PiXi prompts users to explore unusual information right before creating a password, to shake them out of their typical habits and thought processes, and to inspire them to create unique (and therefore stronger) passwords. PiXi's design aims to create an engaging, interactive, and effective nudge to improve secure password creation. We conducted a user study ($N = 238$) to compare the efficacy of PiXi to typical password creation. Our findings indicate that PiXi's nudges do influence users' password choices such that passwords are significantly longer and more secure (less predictable and guessable).

Keywords: Passwords · Authentication · Nudging · User Studies

1 Introduction

Despite decades of development in password authentication alternatives, the majority of websites still require passwords for authentication. Unfortunately, due to time constraints, labor costs, lack of expertise, or apathy, a significant number of people reuse passwords or choose simple, predictable passwords (e.g., birthdays or names). These insecure password choices do not necessarily imply users' lack of intelligence or motivation, but may simply be due to their lack of inspiration or guidance when confronted with a blank password field. Frustration can also arise from unhelpful password policy suggestions, such as "please use special characters to make your password stronger" or "make your password longer to create a strong password." Unfortunately, few solutions exist to support users with creating secure passwords in such helpless situations. While password managers, when used with random password generators, can improve password security [33,32,39], some users are not comfortable using them. Even some official organizations (e.g., governments, enterprises, etc.) do not typically recommend their use for sensitive accounts due to the fear of the password manager vault being compromised. Password manager users still require a strong master password as the key to encrypt the stored passwords in the vault. Therefore, users, regardless of employing password managers or not, still require support for creating secure and memorable

*Contact author: Shengqian Wang, shengqian.wang@ontariotechu.ca

passwords for (at least) these sensitive accounts. Nudging is a promising technique that can encourage users to create more secure and memorable passwords. However, most nudges in password systems apply a one-size-fits-all approach and primarily focus on password meters [25,1,33], which use rigorous password standards to convince users to adjust their passwords to satisfy specific requirements. Unfortunately, many users find effective password meter designs to be annoying [33]. To address these shortcomings, we design ***P**assword **i**nspiration by e**X**ploring **i**nformation (PiXi)*, a novel approach to nudge users towards creating secure passwords. PiXi is the first of its kind that employs a password creation nudge to support users in the task of generating a unique password themselves. PiXi prompts users to explore unusual information right before creating a password, to shake them out of their typical habits and thought processes, and to inspire them to create unique (and therefore stronger) passwords.

We implemented and evaluated a web-based version of PiXi to answer our research questions: (Q1) Which nudges in PiXi are most effective, and do they influence users' password choices? (Q2) Does our PiXi system support users to create more secure passwords? (Q3) How usable is our PiXi system, and how can its usability be improved?

To investigate these research questions, we conducted a user study ($N = 238$) to evaluate the security and usability of passwords generated by users of PiXi. Our contributions and findings include: (i) The design of PiXi—a novel approach to nudging users to create secure passwords. (ii) Security analysis of passwords produced with PiXi. Our study results indicate that PiXi successfully influences users' password choices, such that passwords are longer and more secure (less guessable) than a control group using a typical password creation process. (iii) Usability analysis of the PiXi system. Our study results indicate that PiXi shows promising usability in terms of user perception and memorability. (iv) Analysis of nudge efficacy of PiXi. Our findings indicate that some nudges are more effective than others and that PiXi's combination of nudges do influence users' password choices.

2 Related Work

We first introduce nudging in its most general form, then highlight some of its key applications. We then narrow down our focus to nudges at the time of password creation for graphical passwords and text passwords. Finally, we summarize the key differences between our approach with others.

Nudging. Nudging is a promising strategy to alter people's behavior without limiting their choices or economic incentives [28]. Nudges can successfully change people's decisions by minor and inexpensive interventions [13]. Nudging has been applied in a variety of domains including education [3], ethics [2], social context [20], health [23], finance [6,27], energy savings [11], privacy [1], and security [10]. Computer security experts and administrators have recently been investigating nudges to encourage secure behaviors (see this survey for a great overview [40]).

Password Creation Nudges. Nudging techniques have been employed, with varying degrees of success, to enhance the security of both graphical and text passwords.

Throughout this review, we describe each nudge using the categorizations of Caraban et al. [7].

Nudges in Graphical Passwords. Graphical passwords are a type of knowledge-based authentication that involves remembering (parts of) images instead of a word. Some notable examples of graphical passwords and their variants are Draw-A-Secret (DAS) [19], PassPoints [36], CCP [9], and GeoPass [30,31]), PassFaces [5], and VIP [12]). Background Draw-A-Secret (BDAS)[14] arguably is the first attempt to nudge users away from typical patterns during graphical password selection. It presents users with a background image, on which they need to draw their graphical password. Its background image evokes the "salience bias", thus facilitating the creation of different graphical passwords than if the background image was not present. Zezschwitz et al. [38] used similar nudging techniques to help users create stronger patterns on Android mobile devices. Persuasive Cued Click Points (PCCP) [8] can be considered a facilitate (suggesting alternatives) nudge where users have to select from a point within a randomly positioned view-port (all other options are not available). Some PassPoints variations (e.g., [29,24,21]) can be considered to employ both facilitate (hiding) and reinforce (subliminal priming) nudges. They aim to nudge users away from common patterns by presenting the background image differently at password creation for each user [29,24]. Since these nudges temporarily hide certain options (making them harder to reach), they can be categorized as facilitating (hiding) nudges.

Nudges in Text Passwords. The most straightforward way to nudge strong password selection is to suggest a random password to the user. This is a form of facilitate (default) nudge if implemented so the user has a choice to accept the random password or not. However, memorability is a significant problem for system-assigned random passwords [37]. Password managers can help users remember a random password, but many users still hesitate to adopt them [39]. Even for those users who are successfully nudged to choose a random password and store it in a password manager, it is recommended to avoid using password managers for sensitive accounts [17]. (e.g., email, financial, workplace, etc.) For these reasons, finding other ways to nudge users towards creating secure passwords remains of interest. One way to nudge users towards creating stronger text passwords is through password meters [33]. Employing a confront (friction) nudge, they provide real-time feedback on password strength to motivate users to revise their passwords. Other approaches employ a facilitate (suggesting alternatives) nudge that suggests modifications to the initial password to make it secure [16]. However, these systems are often vulnerable to Guided Brute Force attacks [26].

Our Work vs Others. The existing approaches to nudge stronger text passwords are either (a) default nudges to use a randomly generated password (typically employed as a nudge in password managers [39]), (b) confront (friction) nudges that aim to increase user's awareness of their chosen password's weakness, with no facilitation in coming up with a new password (e.g., password meters [34]), and (c) facilitate (suggesting alternatives) nudges that suggest modifications to a user's initially weak password to make it secure (e.g., [16,22,18]. Our approach with PiXi is entirely different than previous text password nudges; we aim to facilitate the user's password creation without suggesting alternatives, but instead using the following set of nudges immediately prior to

password creation: (i) facilitate (positioning and suggesting alternatives) to help users explore an unusual path (and set of selections) through the PiXi system, (ii) confront (throttling mindless activity) to ensure users consider their PiXi selections, and (iii) reinforce (subliminal priming) to make the user's PiXi selections more prominent and easily accessible at the time the user is attempting to conceive a new password. The goal of this combination of nudges is to create an engaging, interactive, and effective nudge to impact password creation.

3 System Design

The PiXi system aims to nudge users to create stronger passwords, by engaging them with an interactive system for information exploration (e.g., search and select a sequence of keywords) before they create their typical alphanumeric passwords. Instead of limiting user choice, PiXi exposes its users to some unusual and randomized information to shake them out of their typical password creation patterns and get them thinking about new possibilities for their passwords.

PiXi Components. Users interact with PiXi just before password creation through:

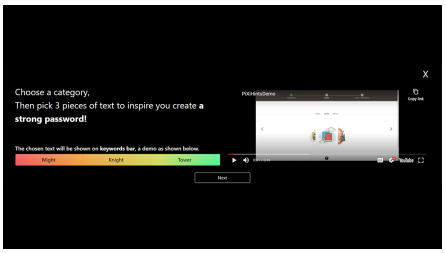
Introduction. The introduction page (see Figure 1a) offers a brief description of the system via a YouTube video tutorial that guides users through the step-by-step process of PiXi. It illustrates how to select a category and a keyword. A short paragraph and a simple animation are also included on the introduction page to assist users in selecting keywords. The users can bypass this page by clicking the "Next" or "X" buttons, and they can always return to it by clicking on the question icon located at the interface.

Category Selection. The category page (see Figure 1b) contains three possible content categories for user selection: images, books, or movies. The order of categories is randomly shuffled for each user. This page contains a *facilitate (positioning) nudge* [7] as it positions a category in the center more prominently to nudge the user to select it. The user still has the option to choose another category. Once a category is selected , the user is directed to an item page (see below).

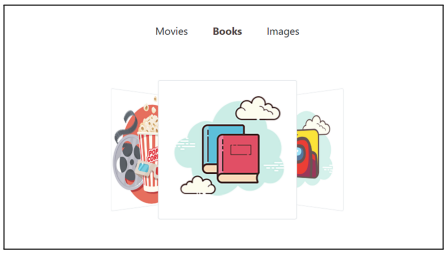
Item Page. The item page contains a set of 20 randomly selected items (e.g., book covers, movie covers, or images) from the selected category.[1] A user then can select an item by clicking its image cover. If not interested in any items, the user can search for her item of interest by the search bar with autocomplete feature. The maximum number of items per page is limited to 20 to maintain an organized user interface. The first row of items, along with the search bar, is shown in Figure 1c. This page contains a *facilitate (suggesting alternatives) nudge* [7], by facilitating the selection from a random set of items over many others.

Keyword Selection Page. After selecting an item, the user is brought to the keyword selection page, where she must choose three keywords. For example, if a user selects "Harry Potter 4" as an item, she will be shown a random excerpt of the book (see Figure

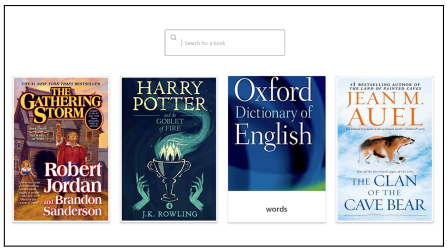
[1] In our PiXi prototype configuration, there are around 6 million possible items (all categories); 20 items are randomly selected from the pool of possible items and shown to the user, for their selected category. However, the number of possible items could be configured to be much larger.



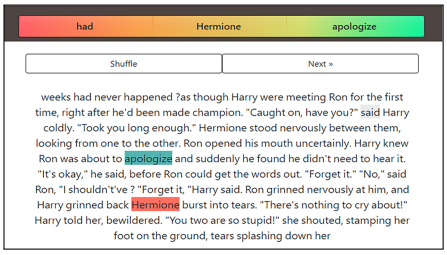
(a) Introduction Page.



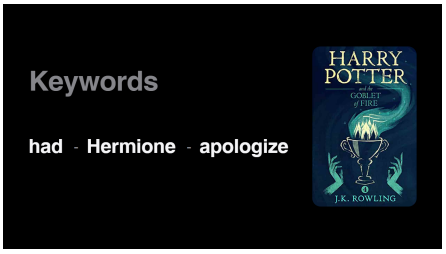
(b) Category Page.



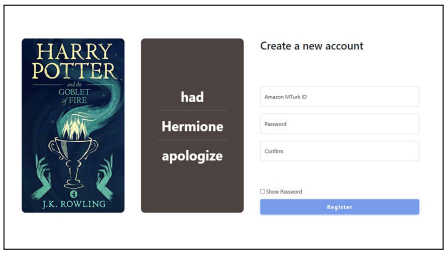
(c) Item Page, Books.



(d) Keyword Selection Page.



(e) Keyword Splash Page.



(f) The Register Page.

Fig. 1: The key user-interaction interfaces in PiXi and its extension PiXi-Hints: (a) the introduction page provides a video tutorial and instructions to users on how to use the system. By clicking the "Next" or "X" buttons, they will be directed to (b) the category page, which contains three possible content categories: Books, Movies, and Images. Once users select their desired category, the user will be taken to (c) the item page, which contains 20 randomly selected items, e.g., book covers in (c). Selecting an item will lead users to (d) the keyword selection page, where they choose three keywords from a random excerpt of the text of the selected item. After selecting all three keywords, users will see the (e) keyword splash page that displays all three chosen keywords (for three seconds) to nudge them further. Finally, users will see (f) the register page which features a large display area of the selected items and keywords on the left side of the typical registration input panel.

1d) from which she is expected to select her keywords. Once each keyword is chosen, it is shown in a bar at the top of the page. We set the maximum number of words per excerpt to 50 to avoid scrolling the page for the user, but the user can click on the shuffle button to land on another random excerpt of the book. After the selection of each keyword, the user is directed to another random excerpt containing the previously selected keyword. Suppose that the user has already selected "had" and "Herminone" as the first and second keywords. For the third keyword selection, she would be shown a random excerpt containing the word "Herminone" (highlighted in red); see Figure 1d for this exact scenario. Then, she can select "apologize" (highlighted in blue once selected) as the final third keyword.

Keyword Splash. Once three keywords are selected, the user will be shown the her selected keywords in a "splash" page as shown in Figure 1e. This page intends to employ further nudging towards selected keywords just before the password creation phase. This page has a black background with soft-white text to create a dramatic color contrast for drawing visual attention to selected keywords, and it automatically close after 3 seconds. But users can manually close it by clicking anywhere on the screen. This splash page aims to offer *a confront nudge* (throttling mindless activity) [7], to nudge users to review the content again.

Registration. PiXi adds a large display area of selected items and keywords on the left side of the typical registration input panel (see Figure 1f). This addition serves *a reinforce nudge (or subliminal priming)* [7], as they make the image cover and keywords more prominent and easily accessible at the time the user is attempting to conceive a new password. We implement the password length requirement of at least 8 characters.

Login. PiXi does not modify the standard login page, and users simply need to enter their username and password to complete the login process.

An Extension: PiXi-Hints. PiXi can also be deployed as a hint for password recovery. To this end, we also have designed a PiXi extension called *PiXi-Hints*, which has all the components of PiXi but slightly differs at the login time. It requires the user to interact with PiXi just before login by inputting their keywords. This interaction intends to help users remember their passwords. In our implementation, we did not require users to recall their keywords but recorded their recall for analysis purposes.[2]

4 User Studies

We conducted a two-session study on Amazon MTurk to evaluate PiXi's ability to nudge users. Our study was approved by our university's Research Ethics Board.

Recruitment and Compensation. Our advertisement was made visible to all MTurk workers, but only US workers with an approval rate of 95% or above were allowed to participate. The users first reviewed and signed the consent form, then were redirected to the PiXi system.

[2] The introduction video had some minor differences for users of PiXi-Hints: they have an additional sentence that advises them to select interesting and memorable keywords. This recommendation is provided to encourage users to remember their keywords as they will need to reuse PiXi to input them again before each login.

	Control	PiXi	PiXi-Hints
Participants	181	185	192
Multi-Identity	76	53	64
Inattentive	15	35	8
Weakly-Committed	19	14	34
Valid Participants (Session 1)	71	83	84
Valid Participants (Session 2)	10	9	12

Table 1: Statistic of session completion and filtered participants across conditions.

The sessions were compensated at the US minimum wage at the time of the study ($7.25/hour). For Sessions 1 and 2 (resp.) with estimated completion times of 7 and 2 minutes (resp.), the participants received $0.85 and $0.35 (resp.).

Conditions/Groups. Upon beginning the study, users were randomly assigned to one of three groups:

1. *Control:* Users create a password and log in as usual (without PiXi).
2. *PiXi:* Users are asked to use PiXi only prior to password creation.
3. *PiXi-Hints:* Users are asked to use PiXi-Hints, which includes using PiXi for both password creation and login.

Sessions and Tasks. Our study contains two sessions. For Session 1, participants were required to register an online account (the process differs based on condition/group), and then complete Questionnaire 1. For Session 2 (7 days later), participants who successfully completed Session 1 were invited back through Amazon MTurk to login to their accounts. After successful login or three unsuccessful login attempts, the participants filled out the exit Questionnaire 2.

Data Cleaning. To maintain the integrity of our analyses, we have rigorously cleaned our data to remove noisy and unreliable instances. For our analyses, Initially, we gathered all the users' responses and stored them in our local database. Next, we proceeded to analyze each record, identifying and quantifying any duplicated data via an automated process. Then, we conducted a manual verification of each record to validate the legitimacy of the users. Afterward, we eliminated the data associated with these participants and kept a copy of the removed entries.: (1) *multi-identity (N=193)*: the users who participated in our study with multiple accounts or bots[3]; and (2) *inattentive (N=58):* users who failed our Likert-scale attention question of "Seven plus three equals eight" in Session 1. (3) *weakly-committed (N=67):* the users with weak predictable passwords (e.g., MTurk IDs or simple number sequences) or inconsistent responses to the SUS scale's Likert questions. Overall, we were surprised by the initial amount of noise in our dataset. The final breakdown of user distribution and removal can be found in Table 1.

Demographics. Table 2 presents an overview of the participant demographics for our study collected through the questionnaire in Session 1. Overall, our participants were

[3] These 193 participants chose an identical but uncommon password, possibly due to these accounts all controlled by one.

composed of 41% female, 58% male, and 1% who preferred not to specify their gender. The majority of participants (51%) fell within the 20–30 age group, followed by the age group of 30–40 making up 32% of participants. Regarding participants' education level, most participants (68%) had a Bachelor's degree, followed by a Master's degree (25%). The majority of participants in our study worked in Business (24%), Technology (21%), or Health (13%).

Gender	**Control**	**PiXi**	**PiXi-Hints**	**Language**	**Control**	**PiXi**	**PiXi-Hints**
Female	42.3%	39.8%	40.5%	English	98.6%	100.0%	98.8%
Male	56.3%	59%	59.5%	Other	1.4%	0.0%	1.2%
N/A	1.4%	1.2%	0.0%	N/A	0.0%	0.0%	0.0%
Age	**Control**	**PiXi**	**PiXi-Hints**	**Occupation**	**Control**	**PiXi**	**PiXi-Hints**
Under 20	0.0%	0.0%	0.0%	Engineering	7.0%	6.0%	7.1%
20-30	54.9%	50.6%	48.8%	Arts and Entmt.	1.4%	4.8%	7.1%
30-40	25.4%	27.7%	34.5%	Business	31.0%	18.1%	26.2%
40-50	11.3%	9.6%	9.5%	Communications	4.2%	2.4%	3.6%
50-60	5.6%	6.0%	6.0%	Social services	5.6%	6.0%	2.4%
60+	2.8%	6.0%	1.2%	Education	7.0%	7.2%	8.3%
N/A	0.0%	0.0%	0.0%	Technology	14.1%	24.1%	23.8%
Education	**Control**	**PiXi**	**PiXi-Hints**	General Labour	2.8%	7.2%	1.2%
None	0.0%	0.0%	0.0%	Agriculture	1.4%	3.6%	3.6%
High School	1.4%	4.8%	8.3%	Government	2.8%	2.4%	2.4%
Bachelor's	74.6%	68.7%	63.1%	Health	18.3%	10.8%	11.9%
Master's	23.9%	24.1%	27.4%	Law	0.0%	0.0%	0.0%
PhD.	0.0%	2.4%	1.2%	Sales	2.8%	4.8%	0.0%
N/A	0.0%	0.0%	0.0%	N/A	1.4%	2.4%	2.4%

Table 2: The user demographics across the three conditions.

5 Results

We begin by evaluating indicators that PiXi's nudges work in Section 5.1. We perform an extensive security analysis in Section 5.2, and usability analysis in Section 5.3.

5.1 Evaluation of Nudging Efficacy

Through various metrics, we evaluate the efficacy of (i) the positioning nudge on the Category Page, (ii) the suggesting alternatives nudge on the Items Page, and (iii) PiXi's overall nudge ability on the users' password.

Positioning Nudge in Category Page. Table 3 shows the acceptance rates of the positioning nudge for categories where one category is initially positioned in the center

	Positioning Nudge (Category Page)	Suggesting Alternatives Nudge (Items Page)
Books	20/56 (35.71%)	40/41 (97.56%)
Movies	29/59 (49.15%)	40/55 (72.73%)
Images	30/51 (58.82%)	63/71 (88.73%)

Table 3: The acceptance rates of the facilitate nudges, combining PiXi and PiXi-Hints.

of the Category Page (for both PiXi and PiXi-Hints). Approximately half of the participants accepted the centered suggested category (especially for Movies and Images). There appears to be a slightly higher preference for the Image category.

Suggesting Alternative Nudge in Items Page. Table 3 also shows the acceptance rates of the suggested alternative nudge in item pages, where the set of 20 randomly selected items initially appeared on the page for both PiXi and PiXi-Hints. Most users (72%-97%, depending on category) accepted one of the suggested items, indicating that this nudge was successful at nudging users towards exploring unique items they might not otherwise consider.

	1 keyword	2 keywords	3 keywords	Total
PiXi	7	12	7	26/83 (31%)
PiXi-Hints	11	14	14	39/84 (46%)
Total	22	26	17	65/167 (39%)

Table 4: The keywords usage rate for both PiXi and PiXi-Hints, including direct and indirect use (e.g., uppercase, lowercase, or additional punctuation added.).

Do PiXi Nudges Influence Resulting Passwords? We aim to determine whether PiXi influenced users' password choices. The most straightforward method to measure this is to determine how many users incorporate their keywords directly in their passwords.

Our findings, shown in Table 4, revealed that 39% of users (31% for PiXi, 46% for PiXi-Hints) incorporated at least one keyword into their passwords. We consider this metric an underestimate of the number of users who are nudged by PiXi, since users may see a relationship between their passwords and keywords that we are unable to detect (e.g., if it is indirectly related and personal in nature). Although it is likely an underestimate, it still provides evidence that a large percentage of users are influenced by the PiXi system during password creation. An emerging critical question is how these nudges have impacted the security of the chosen passwords, which we address next.

5.2 Security Analysis

We study the security of passwords created under each condition from different perspectives including their length, ZXCVBN score, and strength against online and offline attacks. We use a significance level of ($\alpha = 0.05$), and the Holm-Bonferroni correction for multiple-comparison correction. This correction performs an adjustment to significance levels when several statistical tests are performed on a single data set.

Password Length. We recorded the length of the passwords, as one measure of password strength. To determine whether a condition can influence the password length, we test the following Hypothesis:

$\mathcal{H}_0$ *The distribution of password lengths is similar across PiXi, PiXi-Hints, and Control conditions.*

$\mathcal{H}_a$ *The distribution of password lengths differs between PiXi, PiXi-Hints, and Control conditions.*

The one-way ANOVA test ($df = 2$, $N = 238$) rejects the null hypothesis $\mathcal{H}_0$ ($F = 6.5, P = 0.002$) after Holm-Bonferroni correction ($\alpha'_{(1)} = 0.0167$), indicating a significant difference in password length among the three conditions with a large effect size ($\eta^2 = 0.44$). Table 6 shows the mean password length for each condition. The Control condition ($\mu = 9.35$) had a significantly lower password length compared to PiXi ($\mu = 10.87$) and PiXi-Hints ($\mu = 11.42$), while the mean in PiXi and PiXi-Hints are comparable. This suggests that PiXi and PiXi-Hints users tend to create longer passwords than those in the Control condition, which can offer security advantages.

Score	# Guesses X	Description
0	$1 \leq X \leq 10^3$	Too guessable: risky password.
1	$10^3 < X \leq 10^6$	very guessable: protection from throttled online attacks.
2	$10^6 < X \leq 10^8$	somewhat guessable: protection from unthrottled online attacks.
3	$10^8 < X \leq 10^{10}$	safely unguessable: moderate protection from offline slow-hash scenario.
4	$X > 10^{10}$	very unguessable: strong protection from offline slow-hash scenario.

Table 5: ZXCVBN password score range and descriptions [35].

Password Score and Strength. We use ZXCVBN [35], a widely used password meter that is easy to implement and cost-effective. Given an input password, it returns a strength score as described in Table 5. To determine whether a condition can influence the password score, we test the following hypothesis:

$\mathcal{H}_0$ *The distribution of ZXCVBN scores is similar across PiXi, PiXi-Hints, and Control conditions.*

$\mathcal{H}_a$ *The distribution of ZXCVBN scores differs between PiXi, PiXi-Hints, and Control conditions.*

A one-way ANOVA test ($df = 2, N = 238$) revealed a significant difference in password score among the three conditions ($F = 3.868, P = 0.022$) with a medium effect size ($\eta^2 = 0.032$), leading us to reject the null hypothesis $\mathcal{H}_0$ after Holm-Bonferroni correction ($\alpha'_{(3)} = 0.05$). As shown in Table 6, the Control condition with an average of

	Password Length	ZXCVBN Score	SUS Score
Control	9.35 ± 1.73	1.83 ± 1.04	56.60 ± 13.28
PiXi	10.87 ± 4.38	2.16 ± 1.02	54.48 ± 11.93
PiXi-Hints	11.42 ± 4.01	2.31 ± 1.17	56.68 ± 11.49

Table 6: The Mean ± Std. for password length, password score, and SUS score.

($\mu = 1.83$) has a lower password score than PiXi ($\mu = 2.16$) and PiXi-Hints ($\mu = 2.31$). These findings suggest that passwords created through PiXi and PiXi-Hints are stronger than those created by users in the Control condition.

We evaluate password strength by CMU's Password Guessability Service (PGS) [34] which uses numerous state-of-the-art password cracking algorithms to calculate guessability.[4] To assess password strength under online and offline attacks, we employed online and offline attack thresholds of 10^6 and 10^{14} guesses [15]. When a password can be guessed before the online (or offline) attack threshold, we call it *online-unsafe* (or *offline-unsafe*). The summary of our analyses is reported in Table 7. Passwords that can withstand offline attacks in PiXi (14.4%) and PiXi-Hints (32.1%) are significantly higher than in the Control (7%) condition. Conversely, weak passwords are more common in the Control (18.3%) than in PiXi (or 10.8%) and PiXi-Hints (15.5%). We conducted a test to determine whether password strength depends on different conditions, by testing the following hypotheses:

$\mathcal{H}_0$ *The distribution of password strength measurements is similar across PiXi, PiXi-Hints, and Control conditions.*

$\mathcal{H}_a$ *The distribution of password strength measurements differs between PiXi, PiXi-Hints, and Control conditions.*

We performed a χ^2 test ($df = 4, N = 238$) to examine these hypotheses. The results in Table 7 showed a significant difference ($\chi^2 = 17.120$, $P = 0.002$) with a medium effect size ($Cramer's\ V = 0.187$) across different conditions, so we reject the null hypothesis ($\mathcal{H}_0$) after Holm-Bonferroni correction ($\alpha'_{(2)} = 0.025$). This finding further supports that PiXi and PiXi-Hints encourage users to create more unique and stronger passwords than the Control condition.

Should Users Incorporate Keywords in Passwords? As observed in Section 5.1, many users incorporate their keywords into their passwords. Here we aim to determine the security impact of this behavior, to determine whether PiXi should encourage or prevent it. As shown in Table 8, for both PiXi and PiXi-Hints, the passwords using keywords had much higher length, score, and guesses than the average passwords. This suggests that users who used keywords were able to create stronger and longer passwords, and as such future versions of PiXi might encourage this behavior.

Do Some Categories Nudge Stronger Passwords?

[4] We also study them by CKL_PSM–a password strength meter based on the chunk-level PCFG model (CKL_PCFG). However, the results were quantitatively and qualitatively very similar, thus we do not report them here due to space constraints.

	Online-unsafe	Offline-unsafe	Safe
Control	18.3%	74.7%	7%
PiXi	10.8%	74.8%	14.4%
PiXi-Hints	15.5%	52.4%	32.1%

Table 7: Passwords guessability at the online and offline thresholds of 10^6 and 10^{14}, CMU's Password Guessability Service.

	Keywords	Length	Score	CMU Guesses
PiXi	Yes	14.15	2.81	$10^{15.45}$
	No	9.25	1.89	$10^{8.89}$
PiXi-Hints	Yes	13.05	2.51	$10^{14.37}$
	No	9.79	2.17	$10^{10.61}$

Table 8: Comparison of security metrics for passwords with vs. without keywords.

We also investigate whether password strength depends on the nudge category (Books, Movies, or Images). Table 9 shows that passwords created by users who selected Books were most resistant to online and offline attacks. Passwords created by users who selected Images have the least "safe" passwords. One possible reason for this is that keywords from the Images category tend to be less unique compared to the other categories. These results suggest that password strength differs between categories and that future PiXi implementations might avoid using the Images category.

	Online-unsafe	Offline-unsafe	Safe
Books	7.3%	56.1%	36.6%
Movies	14.5%	56.4%	29.1%
Images	15.5%	73.2%	11.3%
Total	14.7%	66.8%	18.5%

Table 9: The guessability of Passwords at the online and offline thresholds across three categories, combining PiXi and PiXi-Hints.

5.3 Usability Analysis

We analyze the usability of PiXi and PiXi-Hints, according to (a) SUS score, (b) user satisfaction, (c) login times, and (d) login rates. Results suggest that PiXi shows promise; most users agreed that it helped them to choose a secure and memorable password, recall rates were promising, and SUS scores were comparable to the Control group.

SUS Score. To measure the usability of the Control, PiXi and PiXi-Hints, we compare the System Usability Scale (SUS)— a commonly-used questionnaire to measure the

SUS Score	Grade	Adjective Rating
> 80.3	A	Excellent
68 – 80.3	B	Good
68	C	Okay
51 – 68	D	Poor
< 51	F	Awful

Table 10: General guideline on the interpretation of SUS score [4].

usability of a system [4]. SUS consists of 10 questions with 5 options to choose from that were asked in our Session 1 questionnaire. The SUS evaluation metrics are shown in 10. As shown in Table 6, the SUS score is very close across conditions, supporting that PiXi has no noticeable usability impact. Although the SUS score is relatively low for PiXi and PiXi-Hints (comparable to Control), this indicates that although PiXi added some steps prior to password creation, that users were not bothered by these steps. As described further below, this may be due to increased user satisfaction that PiXi facilitates creating secure and memorable passwords. To compare PiXi's usability to password meters, where it was found that users were more likely to report creating a password that meets the requirements was difficult [33], we report the relative agreement to Question 8: "The password creation method in this study was easy to use." Our results indicate that PiXi (4.03 ± 0.822) and PiXi-Hints (3.95 ± 0.764) users are more likely to agree the system is easy to use than Control (3.81 ± 0.903), where 1 indicates strong disagreement and 5 strong agreement.

User Satisfaction To determine the extent to which participants value each password system/process, we asked users their level of agreement with the question "I believe this password creating method helped me to choose a secure and memorable text password." Figure 2 gives a visual representation of the distribution of the answers, where 5 is for strongly agree, and 1 for strongly disagree.

The users of PiXi or PiXi-Hints (with averages of 3.95 and 4.05) report higher levels of agreement compared to those in the Control condition (with an average of 2.9). Thus, PiXi and PiXi-Hints systems were successful at inspiring/nudging users to select secure and memorable passwords.

Login Rates and Times.

We analyze our login data from Session 2 for indications of usability and memorability problems in each condition. While the MTurk return rate was low for Session 2, we believe exploring this information can still provide useful insights about system memorability.

Table 11 shows the login success rates (over 3 login attempts) and login time. While the Control group has a higher rate of login failure, we only see this as an indication that PiXi shows promise for helping create stronger and possibly more memorable passwords, and as such further study is required for any concrete statistical analyses.

As shown in Table 11, PiXi-Hints with the additional hint task have higher login times compared to Control. However, surprisingly, PiXi requires a longer login time than Control, while Pixi users tended to require more than one login attempt, which

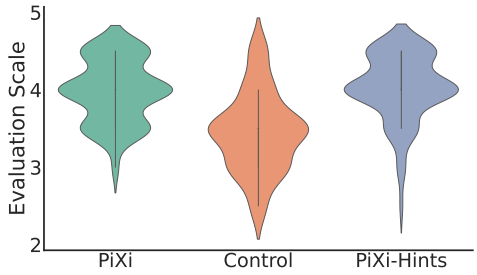


Fig. 2: The violin plot of user satisfaction distributions for three conditions. PiXi and PiXi-Hints users have a similar score distribution, with the majority of users reporting scores of 4 or higher, while Control users have scores concentrated between 3 and 4.

	Control	**PiXi**	**PiXi-Hints**
Login time	14.87 ± 7.38	27.68 ± 22.1	139.5 ± 36.08
Login success rate	7/10 (70%)	8/9 (88.9%)	10/12 (83.3%)

Table 11: Login data for each condition.

increased the average login time. This issue should be analyzed in future work to determine whether it improves over successive logins or not.

6 Conclusion

We designed, implemented, and studied the efficacy of PiXi (**P**assword **i**nspiration by e**X**ploring **i**nformation)—a novel approach to nudge users towards creating secure passwords. PiXi is the first approach we are aware of that employs a text password creation nudge that supports users in the task of coming up with a unique password themselves. PiXi's concept is to ask users to explore unusual information just prior to password creation, to shake users out of their typical habits and thought processes, in the hopes it inspires them to create unique (and therefore stronger) passwords. The results of our study (N = 238) indicate that PiXi is successful at nudging users to create secure passwords, without explicitly asking them to do so. Our findings indicate that PiXi users created passwords that are significantly longer and more resistant to password-guessing attacks. PiXi had a comparable overall perception to typical password creation systems, and users agreed that PiXi helped them to create more secure and memorable passwords. As opposed to password meters, where effective conditions were found to increase difficulty in creating passwords [33], PiXi users found it easier to create passwords.

Our study has some limitations due to Amazon MTurk, which introduced a notable amount of noise in our collected data. While we did our best to fairly catch noise and remove it from our data, it is possible we couldn't catch and filter all noisy data.

However, since the noise should be consistent between each group, any statistically significant finding should be reliable. Future studies should focus on other populations or enhanced methods to filter noise on MTurk. Such future studies should also focus on long-term recall rates and login times over successive logins. It would also be interesting to study whether a shortened version of the PiXi system (e.g., involving only one keyword) could be equally effective at nudging users toward choosing secure passwords.

Future work also includes designing and evaluating extensions to the PiXi system. PiXi presently only offers three categories. In future work, we suggest the study of additional categories (e.g., music/songs, videos, maps, news, or blog posts) to provide users with a broader set of unique paths/nudges through the system.

Our results should stimulate future research into both PiXi itself and more generally novel password creation nudges to support users in secure password creation.

7 Acknowledgment

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

References

1. Acquisti, A., Adjerid, I., Balebako, R., Brandimarte, L., Cranor, L.F., Komanduri, S., Leon, P.G., Sadeh, N., Schaub, F., Sleeper, M., Wang, Y., Wilson, S.: Nudges for privacy and security: Understanding and assisting users' choices online. ACM Computing Surveys (CSUR) **50**(3), 1–41 (2017)
2. Bazerman, M.H., Gino, F.: Behavioral ethics: Toward a deeper understanding of moral judgment and dishonesty. Annual review of law and social science **8**, 85–104 (2012)
3. Breman, A.: Give more tomorrow: Two field experiments on altruism and intertemporal choice. Journal of Public Economics **95**(11-12) (2011)
4. Brooke, J.: SUS: A quick and dirty usability scale. Usability Eval. Ind. **189** (1995)
5. Brostoff, S., Sasse, M.A.: Are passfaces more usable than passwords? a field trial investigation. In: People and computers XIV—usability or else! Springer (2000)
6. Cai, C.W.: Nudging the financial market? a review of the nudge theory. Accounting & Finance **60**(4), 3341–3365 (2020)
7. Caraban, A., Karapanos, E., Gonçalves, D., Campos, P.: 23 ways to nudge: A review of technology-mediated nudging in human-computer interaction (2019)
8. Chiasson, S., Stobert, E., Forget, A., Biddle, R., Van Oorschot, P.C.: Persuasive cued clickpoints: Design, implementation, and evaluation of a knowledge-based authentication mechanism. IEEE Transactions on Dependable and Secure Computing **9**(2), 222–235 (2012)
9. Chiasson, S., Van Oorschot, P.C., Biddle, R.: Graphical password authentication using cued click points. In: ESORICS. vol. 7 (2007)
10. Collier, C.A.: Nudge theory in information systems research a comprehensive systematic review of the literature. Academy of Management Proceedings (1), 18642 (2018)
11. Costa, D.L., Kahn, M.E.: Energy conservation "nudges" and environmentalist ideology: Evidence from a randomized residential electricity field experiment. Journal of the European Economic Association **11**(3), 680–702 (2013)

12. De Angeli, A., Coutts, M., Coventry, L., Johnson, G.I., Cameron, D., Fischer, M.H.: VIP: A visual approach to user authentication. In: Advanced Visual Interfaces (2002)
13. Dijksterhuis, A., Aarts, H., Bargh, J.A., Van Knippenberg, A.: On the relation between associative strength and automatic behavior. Journal of Experimental Social Psychology **36**(5) (2000)
14. Dunphy, P., Yan, J.: Do background images improve "draw a secret" graphical passwords? In: ACM Computer and Communications Security (2007)
15. Florêncio, D., Herley, C., Van Oorschot, P.C.: Pushing on string: The "don't care" region of password strength. Commun. ACM **59**(11), 66–74 (2016)
16. Forget, A., Chiasson, S., van Oorschot, P.C., Biddle, R.: Improving text passwords through persuasion. In: Proceedings of the 4th Symposium on Usable Privacy and Security (2008)
17. Government of Canada: Password managers - get cyber safe. https://www.getcybersafe.gc.ca/en/secure-your-accounts/password-managers#defn-password, online; Accessed: 2023-03-30
18. Houshmand, S., Aggarwal, S.: Building better passwords using probabilistic techniques. In: Annual Computer Security Applications (2012)
19. Jermyn, I., Mayer, A., Monrose, F., Reiter, M.K., Rubin, A.D.: The design and analysis of graphical passwords. In: USENIX Security Symposium (1999)
20. Johnson, E.J., Goldstein, D.: Do defaults save lives? (2003)
21. Katsini, C., Fidas, C., Raptis, G.E., Belk, M., Samaras, G., Avouris, N.: Influences of human cognition and visual behavior on password strength during picture password composition. In: the SIGCHI Conference on Human Factors in Computing Systems (CHI) (2018)
22. MacRae, B.A.: Strategies and Applications for Creating More Memorable Passwords. Master's thesis, Ontario Tech University (2016)
23. Milkman, K.L., Beshears, J., Choi, J.J., Laibson, D., Madrian, B.C.: Using implementation intentions prompts to enhance influenza vaccination rates. Proceedings of the National Academy of Sciences **108**(26), 10415–10420 (2011)
24. Parish, Z., Salehi-Abari, A., Thorpe, J.: A study on priming methods for graphical passwords. Journal of Information Security and Applications **62**, 102913 (2021)
25. Peer, E., Egelman, S., Harbach, M., Malkin, N., Mathur, A., Frik, A.: Nudge me right: Personalizing online security nudges to people's decision-making styles. Computers in Human Behavior **109**, 106347 (2020)
26. Schmidt, D., Jaeger, T.: Pitfalls in the automated strengthening of passwords. In: Annual Computer Security Applications (2013)
27. Thaler, R.H., Benartzi, S.: Save more tomorrow: Using behavioral economics to increase employee saving. Journal of political Economy **112**(S1), 164–187 (2004)
28. Thaler, R.H., Sunstein, C.R.: Nudge: Improving decisions about health, wealth, and happiness (2009)
29. Thorpe, J., Al-Badawi, M., MacRae, B., Salehi-Abari, A.: The presentation effect on graphical passwords. In: the SIGCHI Conference on Human Factors in Computing Systems (CHI) (2014)
30. Thorpe, J., MacRae, B., Salehi-Abari, A.: Usability and security evaluation of GeoPass: A geographic location-password scheme. In: Proceedings of the Symposium on Usable Privacy and Security (2013)
31. Thorpe, J., van Oorschot, P.C.: Human-seeded attacks and exploiting hot-spots in graphical passwords. In: USENIX Security Symposium (2007)
32. Ur, B., Alfieri, F., Aung, M., Bauer, L., Christin, N., Colnago, J., Cranor, L.F., Dixon, H., Emami Naeini, P., Habib, H., Johnson, N., Melicher, W.: Design and evaluation of a data-driven password meter. In: the SIGCHI Conference on Human Factors in Computing Systems (CHI) (2017)

33. Ur, B., Kelley, P.G., Komanduri, S., Lee, J., Maass, M., Mazurek, M.L., Passaro, T., Shay, R., Vidas, T., Bauer, L., Christin, N., Cranor, L.F.: How does your password measure up? The effect of strength meters on password creation. In: USENIX Security Symposium (2012)
34. Ur, B., Segreti, S.M., Bauer, L., Christin, N., Cranor, L.F., Komanduri, S., Kurilova, D., Mazurek, M.L., Melicher, W., Shay, R.: Measuring real-world accuracies and biases in modeling password guessability. In: USENIX Security Symposium (2015)
35. Wheeler, D.L.: ZXCVBN: Low-budget password strength estimation. In: USENIX Security Symposium (2016)
36. Wiedenbeck, S., Waters, J., Birget, J.C., Brodskiy, A., Memon, N.: PassPoints: Design and longitudinal evaluation of a graphical password system. International Journal of Human-Computer Studies **63**, 102–127 (2005)
37. Yan, J., Blackwell, A., Anderson, R., Grant, A.: Password memorability and security: Empirical results. IEEE Security & Privacy **2**(5), 25–31 (2004)
38. von Zezschwitz, E., Eiband, M., Buschek, D., Oberhuber, S., De Luca, A., Alt, F., Hussmann, H.: On quantifying the effective password space of grid-based unlock gestures. In: Mobile and Ubiquitous Multimedia (2016)
39. Zibaei, S., Malapaya, D.R., Mercier, B., Salehi-Abari, A., Thorpe, J.: Do password managers nudge secure (random) passwords? In: Symposium on Usable Privacy and Security (2022)
40. Zimmermann, V., Renaud, K.: The nudge puzzle: Matching nudge interventions to cybersecurity decisions. ACM Trans. Comput.-Hum. Interact. **28**(1) (jan 2021)