Dissecting Nudges in Password Managers: Simple Defaults are Powerful

Samira Zibaei, Amirali Salehi-Abari, and Julie Thorpe, *Ontario Tech University*

https://www.usenix.org/conference/soups2023/presentation/zibaei

This paper is included in the Proceedings of the Nineteenth Symposium on Usable Privacy and Security.

August 7–8, 2023 • Anaheim, CA, USA

978-1-939133-36-6

Open access to the Proceedings of the Nineteenth Symposium on Usable Privacy and Security is sponsored by USENIX.

Dissecting Nudges in Password Managers: Simple Defaults are Powerful

Samira Zibaei
samira.zibaei@ontariotechu.net
Ontario Tech University

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

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

Abstract

Password managers offer a feature to randomly generate a new password for the user. Despite improving account security, randomly generated passwords (RGPs) are underutilized. Many password managers employ *nudges* to encourage users to select a randomly generated password, but the most effective nudge design is unclear. Recent work has suggested that Safari's built-in password manager nudge might be more effective in encouraging RGP adoption than that of other browsers. However, it remains unclear what makes it more effective, and even whether this result can be attributed to Safari's nudge design or simply Safari users. We report on a detailed large-scale study of Chrome users (n=853) aimed at clarifying these issues. Our results support that Safari's nudge design remains more effective than Chrome's among Chrome users. By dissecting the elements of Safari's nudge, we find that its most important element is its *default* nudge. We additionally examine whether a social influence nudge can further enhance the Safari nudge's RGP adoption rate. Finally, we analyze and discuss the importance of a nudge being noticed by users, and its ethical considerations. Our results inform RGP nudge designs in password managers and should also be of interest to practitioners and researchers working on other types of security nudges.

1 Introduction

Passwords remain the most widely-deployed form of authentication over the Web. Users are expected to manage and remember many passwords for many web services and accounts. To cope with remembering numerous passwords, users resort to reusing passwords across different accounts, leaving them vulnerable to credential stuffing attacks [20, 43]. Password managers are one solution to this problem, as long as users make use of their feature to generate and save a random and unique password for each account [30]. Unfortunately, these *randomly generated passwords (RGPs)* are infrequently used (e.g., only 35% of Chrome users [51]). As such, encouraging the use of RGPs in password managers is an important method to improve users' online security [27, 32]. Popular web browsers (e.g., Chrome, Firefox, Safari, etc.) have encouraged the use of RGPs at the time of password creation through *nudging*, without limiting user choices such as typing their own passwords [1–3].[1] A recent study found that Safari users are more likely to adopt an RGP than Chrome or Firefox users [51], suggesting that the underlying cause might be Safari's nudge design. Despite being interesting, this finding has raised many unanswered questions when it comes to the adoption of RGPs: **(Q1)** Does Safari's nudge design remain more effective among users of another browser (e.g., Chrome) or was it just that Safari users were more inclined to adopt RGPs? **(Q2)** Which design components of Safari's nudge (e.g., nudge types it employs) contribute to its high RGP adoption rate? **(Q3)** Can we further extend Safari's RGP adoption rate by incorporating other promising nudging techniques (e.g., social influence[2])?

We evaluate these questions within a population of Chrome users through a large-scale online study (n=853). Participants were asked to register for an account to test a new e-commerce website, during which we observed their interaction with a browser-based password manager. We focus on browser-based password managers to avoid overhead and issues with installing standalone password managers. Each participant

USENIX Symposium on Usable Privacy and Security (SOUPS) 2023.
August 6–8, 2023, Anaheim, CA, United States.

[1] Nudging can broadly be defined as shaping the choice environment to encourage the adoption of a specific choice over others, while not limiting the possible choices [26].

[2] *Social influence nudges*, by providing descriptive information on other people's actions, give the impression that the desired action is approved and accepted by other people [10].

was assigned to one of six nudge design conditions described in Section 3. We perform quantitative analyses on our server logs to find which nudge design works best in terms of RGP adoption rates. To understand users' reasons for adoption (or rejection) of RGPs, we perform a qualitative analysis of users' provided reasons for RGP adoption or rejection. Our key contributions are:

- We provide evidence that Safari's nudge design is indeed more effective than Chrome's. Zibaei et al. [51] only tested Safari's nudge on Safari users; we show the result holds for Chrome users as well.
- We provide the first evaluation of the efficacy of Safari's nudge design elements. Our findings reveal that the default nudge, which automatically populates the password field with a RGP, is the most important element.
- We find that our specific social influence nudge design did not significantly improve the efficacy of Safari's nudge.
- We explore factors that might explain or contribute to the efficacy of nudges in this study. Our findings confirm the results of Zibaei et al. [51], in that they suggest that prior experience using RGPs can have a significant impact on users adopting RGPs.
- We perform a qualitative analysis of users' reasons and barriers to RGP adoption. Surprisingly, security concerns are reasons for and against RGP adoption.

In addition to these contributions, our work sheds light on the importance of whether a nudge is noticed by users. In particular, our results suggest that noticeability is only important for the efficacy of some nudge designs, and that many designs do not take advantage of the attention they draw. We discuss users' reasons for still not adopting the RGP even when they noticed the nudge. Our work provides strong evidence for the effectiveness of default nudges to encourage more secure user behaviors. We further discuss the ethics of default nudges and argue that noticeability can be an important design goal from an ethical perspective.

The remainder of this paper is organized as follows. Section 2 describes related work. The methodology including implementation details and nudge designs is described in Section 3. We report and analyze the results in Section 4. We discuss our findings in Section 5. Concluding remarks and future work are discussed in Section 6.

2 Related Work

Password managers are critical solutions for storing and suggesting secure passwords to users over the Web. However, they are not widely-adopted or at least not used to their full potential [32, 33], partially due to some security and usability concerns. We review the research findings on why password managers are (not) adopted, relevant improvements proposed for password managers, and relevant research on nudging.

2.1 Why Adopt Password Managers?

A growing body of research has focused on understanding which characteristics of password managers and their users contribute to their adoption. Ease of use has been reported as the primary reason for password manager adoption [29]; this can relate to the save password feature [7], auto-fill, user interface design, and ease of installation process [40]. Other reported important features include reliable encryption methods and secure cloud backups of the passwords [28]. While there may be different reasons to adopt password managers for different user groups (e.g., based on gender [49]), in general, it appears that cybersecurity knowledge plays an important role in the adoption of password managers [24, 25].

2.2 Barriers and Improvements

Numerous studies have examined user's barriers to adopting password managers. A variety of obstacles have been identified, including lack of awareness [5, 32, 39], lack of knowledge [5], complex user interface and terminology [5, 40, 41], and lack of trust and transparency [5, 7, 15, 22, 33]. Security concerns such as the risks of a single point of failure [32] and unauthorized access to stored passwords [33] have also hindered adoption. Some users also believe they do not have many accounts to be worried about [32]. Additional reasons for rejecting password managers include the burden of installing standalone software and perceiving them as unnecessary security tools [7, 8, 11].

Efforts to address these adoption barriers have focused on minimizing users' required actions [42], improving password manager user interfaces [9,42], recommending password managers tailored to user requirements [6], introducing educational videos [38], and addressing password reuse issues [41].

2.3 Nudging

Nudging is a behavioral and decision-making technique that influences people's choices without mandating a specific outcome. Nudging theory has been employed in a wide range of human-computer interaction [10] and cybersecurity [52] topics. Some notable examples in cybersecurity and privacy applications involve joining a safe Wi-Fi network [50], making a social post [48], trusting received emails [13], password meters [14, 16], and the problem of password creation in alphanumerical passwords [36, 37, 45, 47] and graphical passwords [31, 44]. Default nudges leverage individuals' tendency to choose the default option instead of considering other alternatives; they have been shown to have a significant effect on changing user behavior across a wide variety of domains [17, 19]. Recently, it has been shown that some popular

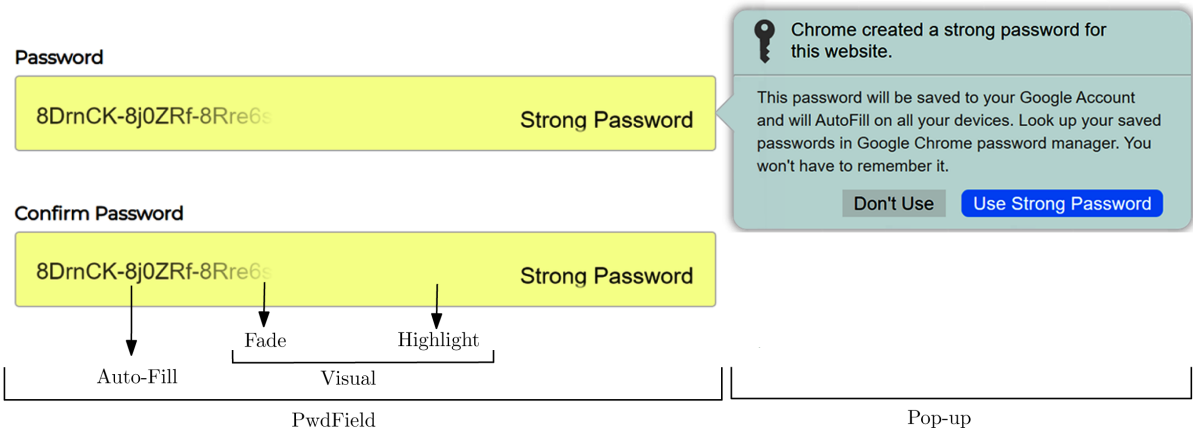

Figure 1: Dissection of Safari's nudge elements: (Left) PwdField is part of the nudge that modifies the password input field, consisting of both Auto-fill and Visual elements. The visual elements are fading the tail of the auto-filled password to make it appear longer and highlighting the password field in yellow. (Right) Pop-up provides a detailed message informing users about the randomly generated password, where it is stored, and reinforcing that they won't need to remember it.

web-browser-based password managers are more effective than others in motivating their users to adopt RGPs, with Safari being the most effective [51].

Our work not only replicates recent findings on password manager nudging [51], but also extends them by answering several critical unanswered questions. Specifically, we investigate whether the high RGP adoption rate of Safari is due to Safari's nudge design or its users. Through a detailed analysis of Safari's UI design elements, we identify and study the specific components that contribute to its high RGP adoption rate. Additionally, we propose and evaluate a social influence nudge enhancement for Safari, which has the potential to further increase RGP adoption rates.

Table 1: Nudge types employed in UI design elements.

UI elements	Default	Social Comparisons	Deceptive Visual	Suggest Alternatives	Just-in-time Prompt
Chrome's UI				✓	✓
Autofill	✓			✓	✓
Pop-up				✓	✓
Visual			✓		✓
Social Pop-up		✓		✓	✓

3 Methodology

Our main goal is to independently evaluate the efficacy of Safari's UI elements in nudging RGP adoption. Our secondary goal is to evaluate the effectiveness of social nudges as an extension of Safari's nudge. To also reproduce findings that suggest Safari is the most effective in nudging RGPs [51], we used a similar methodology and implementation as Zibaei et al. [51][3] but with a population of Chrome users only. We collect quantitative data by collecting user interactions with the password manager, and a post-study questionnaire where participants were asked to answer some questions regarding their actual behavior and intentions. Our study was reviewed and approved by our institution's Research Ethics Board.

[3] We have extended their implementation found at https://github.com/rinoa25/Secure-Password-Creation-Nudge-Prototypes.

3.1 Dissecting Safari's Nudge

We implemented a version of Safari's nudge design (see Figure 1) for Chrome, in order to compare its efficacy to Chrome's nudge between samples of Chrome users. Safari's nudge design can be broken down into various UI elements as labeled in Figure 1. At the highest level, it has two main UI elements: the password field (*PwdField*) and the message box (*Pop-up*). The password field can be further broken down into two UI elements: an auto-filled RGP (*Auto-Fill*) and visual effects (*Visual*). The visual effects *highlight* the password field to make it more visually striking and *fade* the last 6 characters of the suggested password to give the impression of a long password with many characters.

The message box in Safari normally has the heading of *"Safari created a strong password for this website"* and the message of *"This password will be saved to your iCloud Keychain and will AutoFill on all your devices. Look up your saved passwords in Safari Password preferences or by asking*

Table 2: UI design elements for non-Chrome conditions.

Conditions	Safari's UI			Social
	Autofill	Visual	Pop-up	Pop-up
Safari	✓	✓	✓	
PwdField	✓	✓		
PwdField-No-Visual	✓			
Pop-up			✓	
Safari-Social	✓	✓	✓	✓

Siri." This message emphasizes convenience by *"AutoFill"* and explains the storage place by *"iCloud Keychain"*, aiming to educate users on the password manager's functionality. The security is highlighted by a *"Use Strong Password"* button that users must select if they wish to accept the RGP. In our implementation, we have slightly reworded the pop-up message to be consistent with the fact it is running on Chrome (and as such would be saved to the Google Account and can be looked up on the Google Chrome password manager).

The UI elements of Safari as discussed here, as well as the UI of Chrome, employ various types of nudges as described in Table 1 and Section 3.3.

3.2 Prototypes and Conditions

We use a between-groups study design where each group was in a separate condition that used one of the following prototypes. Our prototypes were implemented on the Chrome browser as it is the most popular web browser [4]. Many of the prototypes implement a subset of the UI elements identified in our dissection of Safari's nudge (recall Section 3.1 and Figure 1). The full list of conditions (or prototypes) and their UI elements are summarized in Table 2 and described below:

- The *Safari* prototype simulates Safari's nudge design and interface on Chrome with a minor difference in the wording of the pop-up text to customize it to Chrome: *"Chrome created a strong password for this website. This password will be saved to your Google Account and will AutoFill on all your devices. Look up your saved passwords in Google Chrome password manager. You won't have to remember it"* (see Figure 1). This prototype includes the Autofill, Pop-up, and Visual UI elements.
- The *PwdField* prototype simulates only the PwdField part of Safari's nudge interface. It retains Chrome's informative messaging "Chrome will save this strong password in your Google Account. You won't have to remember it." (see Figure 2). This prototype contains the Autofill and Visual UI elements.
- The *PwdField-No-Visual* prototype simulates only the Autofill UI element of Safari's nudge interface without the visual effects (see Figure 3).

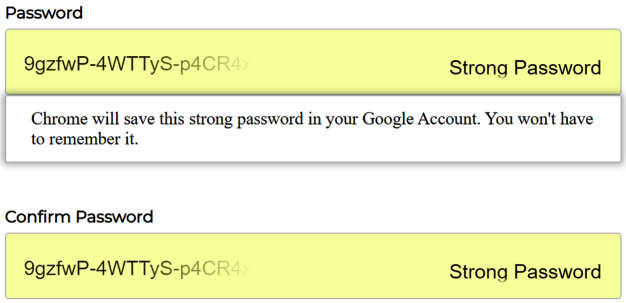

Figure 2: PwdField prototype simulates only the PwdField part of Safari's nudge interface.

- The *Pop-up* prototype simulates only the Pop-up UI element of Safari's nudge interface (see Figure 4).
- The *Safari-Social* prototype (see Figure 5) enhances the Safari prototype with a social influence nudge. This prototype is motivated by the success of social influence nudges in other contexts (e.g., online shopping) [17, 18, 46]. We chose to investigate a social nudge due to its promising results across a variety of fields [17, 18], and also its absence in Safari's nudge design. This prototype includes the Autofill, Pop-up, and Visual UI elements from Safari's nudge, and modifies the Pop-up to include a social nudge in its text and buttons. The Social Pop-up is a pop-up containing the message: "Chrome created a strong password that will be saved and remembered for you", followed by "*Join other users and be part of the secure movement by using this strong password.* [emphasis added] You can look up the saved password in your Google Chrome password manager. You won't have to remember it." The italicized text intends to influence the user by giving the impression that accepting a strong random password is an approved and acceptable action by many other users. We modify the button labels within the pop-up message to reinforce that adopting the RGP is a desirable behavior.
- The *Chrome* prototype simulates Chrome's nudge and interface (see Figure 6). Key phrases of the Chrome popup also exist in our Pop-up prototype: "Use strong password", "You won't have to remember this password", and "it will be saved". Both contain key icons. The main difference is that Chrome's popup contains the RGP within it. Using Caraban's nudging classification [10], both the Pop-up and Chrome prototypes employ facilitate (suggesting alternatives) and reinforce (just-in-time prompt) nudges.

In our study, each prototype discussed above corresponds to a condition. We most often are interested in measuring

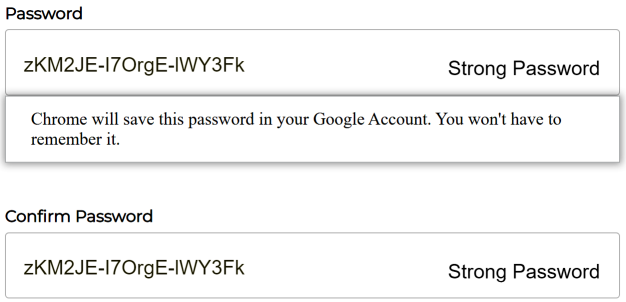

Figure 3: PwdField-No-Visual prototype simulates the PwdField part of Safari's nudge interface, without the Visual effects.

how the RPG adoption rate changes for various conditions with different nudging. We describe how each UI element implements different nudge types next in Section 3.3.

3.3 Nudge Types

We map each UI design element reported in Table 2 to a set of nudge types; see Table 1 for our mapping. Here we describe each of the nudge types and how each design element of Safari employs them:

- *Default* nudges leverage individuals' tendency to choose the default pre-selected option among many other alternatives. By pre-selecting a default option, decision-makers can influence the choices of individuals without restricting their freedom of choice [10]. The Autofill design element employs a default nudge by automatically filling an RGP in the password field.
- *Enabling social comparisons* refers to an individual's tendency to emulate other's behavior. This tendency compels individuals to take heed of the behavior of others and seek social validation when they experience uncertainty in their decision-making [12]. Our Social Pop-up design element implements this through its message "Join other users and be part of the secure movement by using this strong password", intended to evoke a sense that other people are accepting the RGP.
- *Deceptive Visualization* refers to making information relating to desired behaviors more prominent through visual illusions. The goal is to make individuals to focus on a visually-striking option, even if it is not necessarily the best choice [21]. Safari's Visual design element incorporates a fading effect on the characters of the autofilled RGP, creating the illusion that the password is longer than it actually is. Additionally, the Visual design element draws attention to this effect by highlighting the password field in yellow.
- *Suggesting alternatives* bring individuals' attention to the presence of options that may have been overlooked [10]. Except the Visual design element, all other elements in Table 1 employ this nudge type by suggesting the option of selecting an RGP. Chrome provides a small box below the password field that suggests and displays a RGP. Similarly, Autofill automatically fills the password field with a RGP. Pop-up and Social Pop-up suggest using an RGP in a pop-up message. Importantly, none of these design elements force users to choose these alternatives; they are merely presented as options.
- *Just-in-time prompts* seek to grab individuals' attention at the proper time [10]. In the case of encouraging RGP use, the proper time is the time of password creation. Each design element in Table 1 employs a just-in-time nudge by presenting the RGP option immediately after the user clicks on a password field, which is the moment when they are about to create a password.

3.4 Study Structure and Tasks

Participants were randomly assigned to one of the conditions (corresponding to the prototypes discussed in Section 3.2) and were only permitted to complete the study once. Our study, as with Zibaei et al. [51], is structured around six tasks in the following order:

1. *Initial deceptive consent*: To avoid unrealistic, biased user behavior that may draw unrealistic attention to the RGP nudge, we employ a deceptive consent form by deceptively declaring that our study aims at usability testing of an e-commerce website's registration. The users were asked to read and agree to this consent if they wish to continue.
2. *Account registration*: Participants were asked to register on our website, using their email address and password. Participants had the freedom to either create their own password or use a randomly generated password. This is where the nudge design is encountered.
3. *Demographic questionnaire*: Participants were required to answer five demographic questions, where they have the option of "prefer not to answer" for all questions.
4. *Login*: Participants were asked to log in to their accounts using their email address and password.
5. *Post-study questionnaire*: Participants were asked some questions regarding their behavior toward nudges, with an option of "prefer not to answer" for all questions.
6. *Debriefing*: Participants were debriefed about the actual purpose of the study (i.e., nudging) before ending their session, where they were asked to read carefully and agree if they wish to submit their data.

We used the same questionnaires as Zibaei et al. [51], including a post-study attention check question as a means of

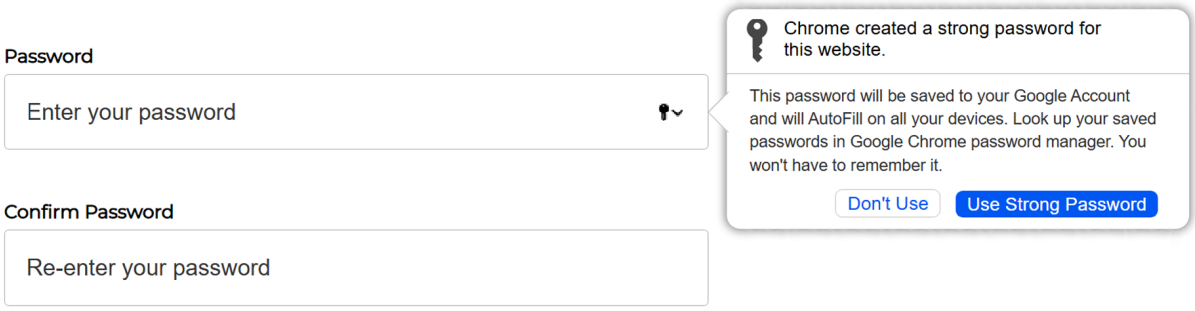

Figure 4: The Pop-up prototype simulates only the Pop-up UI element in Safari's nudge interface.

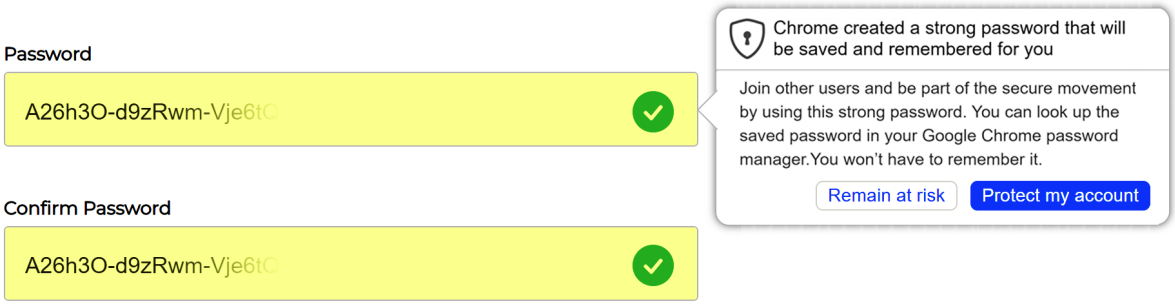

Figure 5: The Safari-Social prototype employs a combination of all nudge types in Table 1. It automatically fills a randomly generated password when the user clicks on the password field, grabs the user's attention by creating the illusion of suggesting a longer password, and suggests a randomly generated password as an alternative with additional social information in a pop-up message.

detecting poor quality data from inattentive participants. In order to minimize potential biases resulting from the questionnaires, we employed recommended guidelines [34].

3.5 Implementation Details

The UIs for each condition are simulated to remotely capture user interactions, while appearing to the user as the browser's PM. In order to achieve this, our website implements the UIs, and we don't define the password field as a proper password input to prevent the browser's PM from being invoked/interfering. The users were not aware of this simulation until the end of the study when we revealed it in a debriefing task. We record user interactions with the simulated password manager while creating an account and logging into the website. To ensure the confidentiality and quality of our collected data, we take a few key measures: (a) we ensure the actual built-in password manager of Chrome is not invoked for the account creation and login process; (b) we only collect passwords with anonymous identifiers, and we do not collect email addresses; (c) we only collect data once users have submitted the final debriefing form; and (d) users were allowed to participate only once in our study. As the participants of our study were expected to use Chrome as a web browser, we employed the user-agent header to confirm the correct browser was in use.

3.6 Recruitment

For our study, we recruited 862 participants via Amazon's Mechanical Turk platform, restricted to individuals from the United States. The duration to complete all study tasks was estimated to be less than 5 minutes. In accordance with the minimum wage in the United States ($7.25 USD per hour), participants were compensated $0.60 USD for their participation.

3.7 Analysis Method

We conduct both quantitative and qualitative analyses of our collected data. Our analyses aim to determine the effectiveness of Safari's nudge design and identify the key nudging elements that contribute to its effectiveness in encouraging RPG adoption. Specifically, we seek to determine whether there are statistically significant differences in RPG adoption

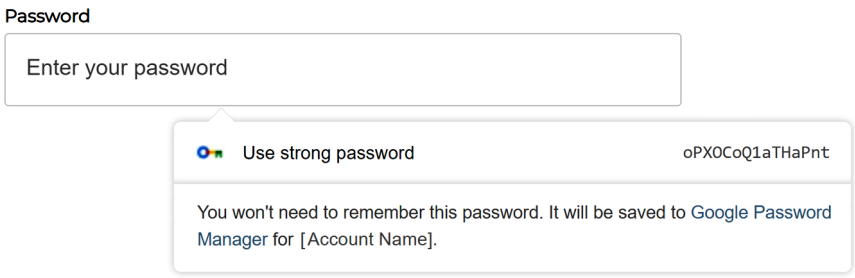

Figure 6: Chrome prototype simulates Chrome's nudge and interface (since Chrome 105, released Aug. 2022).

rates across the various condition groups, where all users in each group are exposed to a specific prototype introduced in Section 3.2. We use the χ^2 test, a statistical method for comparing proportions, to determine the statistical significance of adoption rates. For two condition groups of A and B (e.g., Safari vs. Chrome), our null and alternative hypotheses are:

H_0 The randomly generated password adoption rates are similar for two condition groups A and B.

H_a The randomly generated password adoption rates differ between two condition groups A and B.

We also use Bonferroni multi-test correction, and we report a result as statistically significant when it is significant after correction.

For the qualitative analysis, we assessed participants' comments in a post-study questionnaire, including an open-ended question that asked the participants to explain their reasons for (not) using an RGP. These analyses allow us to gain a deeper understanding of the participants' rationale for or against the adoption of randomly generated passwords. Using the codes reported in Zibaei et al. [51], two researchers in our lab independently categorized users' comments and their rationale. Participants who provided multiple reasons for their password creation behavior were given multiple codes. To ensure the consistency of our coding process, we used Cohen's Kappa to measure inter-rater agreement. The Cohen's Kappa score was $\kappa = 0.95$, demonstrating "almost perfect agreement" between the two coders.

4 Results

Here we describe our participant demographics, discuss the efficacy of Safari's nudge elements and our enhancements, and present factors and reasons behind adoption and abandonment behavior.

4.1 Participant Demographics

Table 9 presents an overview of the demographics of our participants across all conditions. A total number of 896 participants initially signed up for our study, where 34 participants opted out before debriefing. From the remaining 862 participants, 9 respondents were removed due to being duplicates or failing in answering the attention question. Some notable statistics of our study across all conditions (not reported in Table 9) follow. Participants identified as 58.7% male, 40.7% female, and 0.6% preferred not to answer. Participants were mostly 26-35 years old. Of their education, 68.1% had a bachelor's degree and almost half of our participants had a business and IT background.

4.2 Safari's Design or its Users? (RQ1)

Zibaei et al. [51] only evaluated Safari's nudge among Safari users. This motivates us to ask whether Safari's nudge design remains more effective for users of another browser? In particular, among Chrome users, does Safari's nudge design remain more effective than Chrome's? We compare the RGP adoption rates between condition groups Chrome and Safari (67.6% vs 81.1%). Using the χ^2 test, we reject the null hypothesis ($\chi^2 = 7.95$, $p = 0.0038$) with small effect size (Cramer's $V = 0.16$). See Table 3 for more details. This result implies that Chrome users are more likely to adopt a randomly generated password when exposed to Safari's nudges compared to Chrome's nudges, supporting that the effectiveness of Safari is likely due to its design.

4.3 Which Nudge Elements? (RQ2)

Which elements of Safari's nudge contribute to its high RGP adoption rate? To answer this question, we first attempt to understand whether the nudges involved in PwdField are more effective than those of Pop-up (see Figure 1). So, we compare RGP adoption rates between condition groups PwdField and Pop-up (75.2% vs 57.9%). Using the χ^2 test, we reject the null hypothesis ($\chi^2 = 9.33$, $p = 0.0022$) and the effect size is weak (Cramer's $V = 0.18$). Thus, users are more likely to use a randomly generated password when they are exposed to PwdField (which employs default, deceptive visualization, suggesting alternative, and just-in-time prompt

Table 3: χ^2 test results indicate that Safari's nudge is significantly more effective than Chrome's even for Chrome users, and PwdField is significantly more effective than Pop-up. PwdField-No-Visual and PwdField are statistically comparable in their nudging ability. The Safari-Social nudge offered no significant improvements over Safari's nudge. *Significance level $\alpha = 0.005$.

Research Question	Test Description	df	N	χ^2	p	V
RQ1	Chrome vs. Safari	1	282	7.95	0.0038*	0.16
RQ2	PwdField vs. Pop-up	1	279	9.33	0.0022*	0.18
	PwdField vs. Safari	1	296	2.386	0.122	N/A
	Safari vs. Pop-up	1	269	19.657	< 0.00001*	0.27
	Chrome vs. PwdField-No-Visual	1	281	9.077	0.002*	0.17
	Chrome vs. Pop-up	1	265	2.66	0.102	N/A
	PwdField vs. PwdField-No-Visual	1	295	2.79	0.09	N/A
RQ3	Safari-Social vs. Safari	1	293	0.58	0.80	N/A

Table 4: RGP adoption rate across prototypes

	PwdField-No-Visual	**Safari**	**Safari-Social**	**PwdField**	**Chrome**	**Pop-up**
RGP adoption rate	83.1%	81.1%	80.0%	75.2%	67.7%	57.9%

nudges) compared to Pop-up (which employs only suggesting alternative and just-in-time prompt nudges).

One might wonder if the deceptive visualization nudge in visual effect (i.e., Highlight and Fade) increases the RGP adoption rates for PwdField. To test this, we compare the RGP adoption rates between condition groups of PwdField and PwdField-No-Visual (75.9% vs 83.1%), which don't exhibit a statistically significant difference ($\chi^2 = 2.79$, $p = 0.09$). Thus, the Visual element doesn't appear to improve RGP adoption rates over just the Autofill element alone (one can even note the higher adoption rate of 83.1% for PwdField-No-Visual). Table 3 highlights interesting findings obtained through pairwise comparisons between conditions. Taken together, these results indicate that Autofill is the most important element of Safari's nudge. Autofill is the only part of the interface that implements a *default nudge*. It provides evidence that using a simple default is the most effective type of nudge for encouraging RGP adoption. However, other UI elements might have other advantages as discussed in Section 5.

4.4 Can a Social Nudge Improve? (RQ3)

Can we further extend Safari's RGP adoption rate by incorporating a social influence nudge? We compare RGP adoption rates between condition groups Safari and Safari-Social (81.1% vs 80.0%). The null hypothesis holds true ($\chi^2 = 0.58$, $p = 0.8$). Thus, users exhibit a similar likelihood of selecting RGPs when they are exposed to Safari and Safari-Social. We conclude that our prototype, which was a nudge to enable social comparisons, could not enhance the Safari nudge's ability to encourage more users to adopt RGPs.

4.5 Contributing External Factors

Do the external factors identified in other studies [51] (i.e., nudge noticeability and previous experience using RGPs) impact RGP adoption rates in our study? We compare RGP adoption rates of the users who self-identify as having previously used RGPs and who had not (78.8% vs. 20.0%). We reject the null hypothesis ($\chi^2 = 37.44$, $p < 0.0001$) with small effect size (Cramer's $V = 0.20$). This suggests that prior experience with randomly generated passwords influences their adoption and usage.

We also compare RGP adoption rates of the users who answered yes to the same post-study question as other work [51]: "Did you notice the recommendation to use a random password while registering on our website?" and those who reported had not (63.7% vs. 2.87%). We reject the null hypothesis ($\chi^2 = 26.97$, $p < 0.0001$), with small effect size (Cramer's $V = 0.17$). This suggests that participants who noticed the nudge were more likely to adopt a RGP compared to the participants who did not notice the nudge. We examine the issue of noticeability in more detail, for each prototype, in Section 5.

4.6 RGP Adoption Reasons and Barriers

Why do people adopt or reject RGPs? To gain insight, we administered a post-study questionnaire to ask why they either selected or rejected the RGP by asking "Can you describe the reason why you used/did not use the random password generator?". The results (see Tables 5 and 6) revealed that the main reason for adopting an RGP is the security it offers (35.76% of total participants). Convenience is the second most common adoption reason (13.72% of total participants). Interestingly, security concerns are the primary barrier to

RGP adoption (8.91% of total participants). Most security concerns refer to password manager's potential vulnerabilities related to password vault breaches, and privacy and safety issues. The second most common rejection reason is the issue of memorability. Participants preferred selecting memorable passwords to ease their use across multiple devices. However, this perception of being unable to use password managers across multiple devices may stem from a lack of knowledge regarding the functionality of password managers. It may also be due to not using the same browser across multiple devices, or simply not trusting them to sync.

5 Discussion

We discuss the interpretations of our findings, some more exploratory findings, as well as other considerations of interest. In particular, we examine the interplay between RGP adoption and users noticing the nudge in Section 5.2. We discuss the ethics of default nudges and value of other nudge types in Section 5.3. We end this section with a discussion of limitations in Section 5.4.

5.1 The Power of Simple Default Nudges

By dissecting Safari's nudge and examining its element's efficacy in nudging, we find that Autofill is the most powerful at encouraging RGP use. Autofill is implementing a simple default nudge by automatically filling in a suggested RGP for the user. This simple default creates some friction for users who wish to choose their own password, since they need to either navigate to the "Don't Use" button or manually delete the RGP. It also prominently reinforces that keeping the RGP is the recommended action.

A closer examination of the rates of RGP adoption across all conditions also shows that the prototypes incorporating the default nudge (Safari, Safari-Social, PwdField, and PwdField-No-Visual) have RGP adoption rates that are 75-83%, whereas the prototypes that do not use a default nudge (Chrome and Pop-up) have RGP adoption rates between 58-68%. Grouping all data from conditions that incorporate a default nudge vs. those conditions that do not incorporate default nudges, we find the presence of a default nudge is more effective at encouraging RGP adoption ($\chi^2 = 28.27$, $p < 0.0001^*$, $V = 0.18$).

Our work supports that default nudges are quite powerful at encouraging RGP use. Our findings are in line with reviews that found default nudges are one of the most effective types of nudge across different domains and applications [17]. We believe this is good news for deployment of security nudges in general, as default nudges are easy to implement, and have less parameters to adjust in the design that can lead to its success or failure. Even something as simple as a pop-up that intends to suggest alternatives can fail due to subtle choices in words, colors, positioning, etc. Visualizations can be even more challenging to design. However, such elements may improve default nudge designs from an ethical perspective (see Section 5.3 for further discussion).

5.2 Is Noticing the Nudge Important?

We explore whether *noticeability* (i.e., how noticeable a nudge is) might be a cause of some prototypes being more effective than others. For each prototype, Table 7 shows the relationship between noticing the nudge and RGP adoption. While we found in Section 4.5 that participants (across all conditions) who noticed the nudge were more likely to adopt an RGP, a closer inspection using Pearson correlation reveals that RGP adoption is only positively correlated with noticing the nudge in the Chrome and Safari prototypes. All other conditions had no noticeable correlation and in one prototype (PwdField), negative correlation. The positive correlation only being in the Chrome and Safari groups implies that familiarity with the interface can lead to higher trust and subsequent RGP adoption. A closer examination of the other prototype nudges (which are all novel to Chrome users, since they are neither exactly the Chrome or Safari nudge) reveals some interesting insights. In particular, Pop-up was comparably noticeable to the other conditions, yet had a significantly lower rate of adoption; it was the only novel prototype not involving a default nudge. Another interesting comparison point is PwdField vs. PwdField-No-Visual. When the visual element was missing, more participants chose to adopt the RGP (both in the group that noticed the RGP and who did not notice the RGP). One possible reason is the interface drawing less attention to itself and therefore caused fewer participants to hesitate and seriously weigh their options. We discuss this issue further in Section 5.3.

We analyzed user's comments to gain a deeper understanding of why some individuals chose to reject the RGP even after noticing the nudge. The most common barrier among participants who acknowledged the nudge but rejected the RGP was security concerns (22.75% of the participants). The second most commonly mentioned barrier was the difficulty of memorizing the RGP (17.96% of participants). These reasons (and their percentages) are comparable to all users who rejected the RGP, regardless of whether they noticed the nudge, so it appears that noticing the nudge is simply not enough to change some users' beliefs about the security offered by RGPs and usability of password managers.

5.3 Ethics of Default Nudges

We were surprised to observe that PwdField-No-Visual was more effective than PwdField—we had expected that the Visual element of PwdField would make the prototype more visually striking, leading to higher rates of RGP adoption. One possible explanation is that by the interface drawing less attention to itself, fewer participants took enough notice to seriously consider the implications of adopting the RGP.

Table 5: Reasons for Adopting the RGP

Code	N	%	Sample of Comments
Security	305	35.7%	*"I used it because it gave me a strong password"*
Convenience	117	13.7%	*"It was easier than coming up with my own password."*
Noise	100	11.7%	*"Z9tOh\ES*GOX"*
User preference	49	5.7%	*"I always use a random password generator."*
Incongruous	48	5.6%	*"I'd rather make my own"*
Remember password feature	38	4.4%	*"I always do them when I can and save it to my computer/Google for ease of login"*
Didn't care about the website	11	1.3%	*"I thought I should because I am doing a HIT."*
Unsure	8	0.9%	*"I don't know how to use it and have never really heard of it until now."*
Strict password policy	3	0.3%	*"Use random password generator because we can't match the requirement."*

Table 6: Reasons for Rejecting the RGP

Code	N	%	Sample of comments
Security concern	45	5.8%	*"The random password generator wasn't running locally on the CPU; so it was insecure."*
Memorability issue	43	5.0%	*"I would rather use a password that I can memorize.*
Noise	39	4.7%	*"None"*
Incongruous	35	4.1%	*"Using the random password is very difficult to hack"*
Trust issue	30	3.5%	*"I did not trust it"*
User preference	26	3.0%	*"I can create my own password"*
Didn't care about the website	8	0.9%	*"Not sure if I will keep this account."*
Didn't notice the nudge	6	0.7%	*"I did not see that option"*
The desire to reuse password	4	0.4%	*"I like to use similar passwords for each website."*
Lack of knowledge	2	0.2%	*"I was unaware of it."*

This, combined with the observed higher efficacy rates of the default nudges, raises the question of whether default nudges have ethical considerations? What if the user doesn't stop to consider the implications of accepting the default? If the user fails to notice that a default has been set, which they have a choice to accept or reject, then has something unethical occurred (even if it is the "best choice")?

From Table 7, we notice that for most conditions, the percent of users who didn't notice the nudge and rejected the RGP is higher than the percentage of users who noticed the nudge and rejected the RGP. For the PwdField and PwdField-No-Visual groups, this effect is reversed with most of the users who didn't notice the nudge accepting the RGP. These are the two groups that employ default nudges but not pop-up messages. This observation raises the question of whether pop-up messages draw a user's attention that they need to make a decision. Given these observations, Pop-up and Visual elements, while not more effective at encouraging RGP adoption, have advantages from an ethical perspective. We believe further research on nudging should also consider additional metrics of success. For example, in the context of RGP nudges, perhaps to consider the user's understanding of the decision they made and its implications.

5.4 Limitations

We caution readers against interpreting our raw percentages as rates of RGP adoption in other non-experimental settings. This is due to the prevalence of low-quality data from the Amazon MTurk platform [23]. However, the comparisons between the different conditions we test should have validity, as the amount of low-quality data (or noise) should be similar between each of the conditions we test. To reduce the impact of poor data quality, we add a question that aims to catch inattentive participants, which we have excluded from our analysis.

For all non-Chrome conditions (e.g., Safari, PwdField, etc.), Chrome users might have noticed the interface was different than usual. One might ask whether this could explain Safari's higher RGP adoption rate, since this novelty might have brought additional salience to the nudge. To this end, we examine the relationship between noticing the nudge and RGP adoption in Section 5.2. Our findings in that section indicate that Chrome and Safari conditions have similar noticeability ($\chi^2 = 3.59$, $p = 0.057$) and some Chrome users (20.1%) rejected the RGP despite noticing the nudge. Thus, appears that Safari's higher success rate is not because of being more noticeable, but in nudging users towards accepting the RGP after gaining their attention. To further reflect on whether novelty could explain our result, we discuss the success rates of our other prototypes that should be novel to Chrome users.

Table 7: For each prototype, the relationship between noticing the nudge (Noticed Y/N) and adopting the RGP (Y/N). Percentages are shown as well as Pearson correlation values (r).

		Chrome		Safari		PwdField		PwdField-No-Visual		Pop-up		Safari-Social	
RGP		Y	N	Y	N	Y	N	Y	N	Y	N	Y	N
Noticed	Y	63.76%	20.1%	77.6%	13.28%	65.35%	22.78%	77.46%	12.67%	52.38%	34.12%	76.66%	16.66%
	N	2.87%	10.79%	4.19%	4.19%	9.80%	1.96%	5.63%	2.81%	5.55%	7.14%	3.33%	2.66%
Cor. r		0.4		0.24		-0.69		0.14		0.11		0.16	

Table 8: For each prototype, the percent of users who reported: accepting the RGP due to beliefs it was secure (Security), and rejecting the RGP due to trust or security concerns (Mistrust).

	Chrome		Safari		PwdField		PwdField-No-Visual		Pop-up		Safari-Social	
RGP	Y	N	Y	N	Y	N	Y	N	Y	N	Y	N
Security	48.95%	N/A	46.61%	N/A	42.06%	N/A	44.27%	N/A	54.66%	N/A	37.50%	N/A
Mistrust	N/A	32.60%	N/A	13.63%	N/A	12.5%	N/A	25.00%	N/A	13.20%	N/A	44.00%

We observe that the Pop-up prototype, despite being a novel interface, did not result in higher RGP adoption than Chrome (in fact it was nearly 10% lower).

One might wonder if novelty has introduced a lack of trust in the interface. However, the number of participants who cited trust or security concerns as reasons for rejecting the RGP are overall quite low (see Table 8). Notably, the percentage of participants who cited mistrust was higher in Chrome than in the other conditions (except Safari-Social), indicating that this issue was not prevalent among non-Chrome prototypes/conditions. Despite these observations, it remains possible that novelty may have somewhat increased Safari's nudge success. However, our comparisons between the elements of Safari's nudge interface should be equally impacted by novelty.

We implemented a particular social nudge design in our study, which shows no significant effect toward further improving Safari's nudge design (in terms of RGP adoption). Our result does not suggest the inefficiency of social nudges in general, but rather indicates that our specific design did not produce the desired effect.

Our study is limited to the evaluation of password nudges solely within the context of web browsers; it is possible that results may differ between desktop and mobile devices. It is also worth noting that the wording of the messaging in Safari and Chrome's UI has changed since our study was conducted.

Our study is conducted with a limited diversity of participants on the Amazon MTurk platform, where all of our participants were from the United States and are fluent in English, which may have resulted in a language or cultural bias. While it is shown that MTurk workers are more tech-savvy and younger, previous research implies that online privacy and security behavior studies can still estimate the general population's behavior [35]. However, MTurk users may encounter more account creation scenarios than the general population, leading to a higher rate of RGP adoption.

We collect data on users' behavior when signing up once to test the usability of the registration page. However, users' behavior might differ when the user signs up with the intention of long-term use. Further study is needed to determine whether planned long-term use might reduce RGP adoption rates.

Some users may have used other methods to generate random passwords; as such, we record entered passwords. Our analysis reveals that 8.1% of users demonstrated such behavior, with the following breakdown: Safari (2.8%), Chrome (10.8%), Pop-up (7.9%), PwdField (9.2%), PwdField-No-Visual (8.5%), and Safari-Social (9.3%).

6 Conclusion

Our work provides clarity on a number of issues brought up in other research on password manager RGP nudges. In particular, we offer evidence that Safari's password manager nudge is more effective at encouraging RGP adoption than Chrome's. Additionally, we find which nudge types are most effective at encouraging RGP use—it turns out that simple default nudges are the most powerful. While the other nudges we studied were less effective at encouraging RGP use, they may still serve an ethical purpose in increasing awareness to users regarding their decision.

Future work includes addressing the "missed opportunity" observed in many of the nudge prototypes we studied, where the nudge was noticed but unfortunately failed to capitalize on the user's attention. Future attempts at improving these nudge designs should focus on addressing the main barriers

we identified: concerns about security/trust and the possibility of needing to remember the RGP. Educating users about how password managers work might help. Future work also includes developing approaches to personalize these security nudges to improve their efficacy.

Acknowledgments

This research was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC). We thank Nicholas Hughes for his assistance with our study.

References

[1] Autofill your user name and password in Safari on mac. https://support.apple.com/en-ca/guide/safari/ibrwf71ba236/mac. Accessed: 2023-02-10.

[2] How to generate a secure password in Firefox. https://support.mozilla.org/en-US/kb/how-generate-secure-password-firefox. Accessed: 2023-02-10.

[3] Let Chrome create and save a strong password for your online accounts. https://support.google.com/chrome/answer/7570435?hl=en&co=GENIE.Platform%3DDesktop. Accessed: 2023-02-10.

[4] Market share held by leading desktop internet browsers in the Uunited States from January 2015 to August 2022. https://www.statista.com/statistics/272697/market-share-desktop-internet-browser-usa/. Accessed: 2023-02-15.

[5] Nora Alkaldi and Karen Renaud. Why do people adopt, or reject, smartphone password managers? In *European Workshop on Usable Security*, 2016.

[6] Nora Alkaldi and Karen Renaud. Encouraging password manager adoption by meeting adopter self-determination needs. In *Proceedings of the 52nd Hawaii International Conference on System Sciences*, 2019.

[7] Fahad Alodhyani, George Theodorakopoulos, and Philipp Reinecke. Password managers—it's all about trust and transparency. *Future Internet*, 12(11):189, 2020.

[8] Sal Aurigemma, Thomas Mattson, and Lori Leonard. So much promise, so little use: What is stopping home end-users from using password manager applications? In *Hawaii International Conference on System Sciences*, 2017.

[9] Jannatul Bake Billa, Anika Nawar, Md Maruf Hasan Shakil, and Amit Kumar Das. Passman: A new approach of password generation and management without storing. In *IEEE International Conference on Smart Computing & Communications (ICSCC)*, pages 1–5, 2019.

[10] Ana Caraban, Evangelos Karapanos, Daniel Gonçalves, and Pedro Campos. 23 ways to nudge: A review of technology-mediated nudging in human-computer interaction. In *ACM Conference on Human Factors in Computing Systems (CHI)*, pages 1–15, 2019.

[11] Sonia Chiasson, Paul C van Oorschot, and Robert Biddle. A usability study and critique of two password managers. In *USENIX Security*, 2006.

[12] Robert B Cialdini and N Garde. Influence (vol. 3). *Port Harcourt: A. Michel*, 1987.

[13] Molly Cooper, Yair Levy, Ling Wang, and Laurie Dringus. Subject matter experts' feedback on a prototype development of an audio, visual, and haptic phishing email alert system. *Online Journal of Applied Knowledge Management*, 8(2):107–121, 2020.

[14] Serge Egelman, Andreas Sotirakopoulos, Ildar Muslukhov, Konstantin Beznosov, and Cormac Herley. Does my password go up to eleven? the impact of password meters on password selection. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 2379–2388, 2013.

[15] Michael Fagan, Yusuf Albayram, Mohammad Maifi Hasan Khan, and Ross Buck. An investigation into users' considerations towards using password managers. *Human-centric Computing and Information Sciences*, 7(1):1–20, 2017.

[16] Maximilian Golla, Björn Hahn, Karsten Meyer zu Selhausen, Henry Hosseini, and Markus Dürmuth. Bars, badges, and high scores: On the impact of password strength visualizations.

[17] Dennis Hummel and Alexander Maedche. How effective is nudging? a quantitative review on the effect sizes and limits of empirical nudging studies. *Journal of Behavioral and Experimental Economics*, 80, 2019.

[18] Moritz Ingendahl, Dennis Hummel, Alexander Maedche, and Tobias Vogel. Who can be nudged? examining nudging effectiveness in the context of need for cognition and need for uniqueness. *Journal of Consumer Behaviour*, 20(2):324–336, 2021.

[19] Jon M Jachimowicz, Shannon Duncan, Elke U Weber, and Eric J Johnson. When and why defaults influence decisions: A meta-analysis of default effects. *Behavioural Public Policy*, 3(2):159–186, 2019.

[20] David Jaeger, Chris Pelchen, Hendrick Graupner, Feng Cheng, and Christoph Meinel. Analysis of publicly leaked credentials and the long story of password (re-) use. *Hasso Plattner Institute, Universidad de Potsdam. Disponible en https://bit. ly/2E7ZT01*, 2016.

[21] Daniel Kahneman, Stewart Paul Slovic, Paul Slovic, and Amos Tversky. *Judgment under uncertainty: Heuristics and biases*. Cambridge university press, 1982.

[22] Ambarish Karole, Nitesh Saxena, and Nicolas Christin. A comparative usability evaluation of traditional password managers. In *International Conference on Information Security and Cryptology*, pages 233–251. Springer, 2010.

[23] Ryan Kennedy, Scott Clifford, Tyler Burleigh, Philip D Waggoner, Ryan Jewell, and Nicholas JG Winter. The shape of and solutions to the mturk quality crisis. *Political Science Research and Methods*, 8(4):614–629, 2020.

[24] Shelia M Kennison and D Eric Chan-Tin. Predicting the adoption of password managers: A tale of two samples. *TMS Proceedings 2021*, 2021.

[25] Shelia M Kennison, Ian T Jones, Victoria H Spooner, and D Eric Chan-Tin. Who creates strong passwords when nudging fails. *Computers in Human Behavior Reports*, 4:100132, 2021.

[26] Thomas C. Leonard, Richard H. Thaler, and Cass R. Sunstein. Nudge: Improving decisions about health, wealth, and happiness, 2008.

[27] Michael D Leonhard and VN Venkatakrishnan. A comparative study of three random password generators. In *IEEE International Conference on Electro/Information Technology*, pages 227–232, 2007.

[28] Raymond Maclean and Jacques Ophoff. Determining key factors that lead to the adoption of password managers. In *IEEE International Conference on Intelligent and Innovative Computing Applications (ICONIC)*, pages 1–7, 2018.

[29] Peter Mayer, Collins W Munyendo, Michelle L Mazurek, and Adam J Aviv. Why users (don't) use password managers at a large educational institution. In *USENIX SOUPS*, 2022.

[30] Sean Oesch and Scott Ruoti. That was then, this is now: A security evaluation of password generation, storage, and autofill in browser-based password managers. In *USENIX Security Symposium*, 2020.

[31] Zach Parish, Amirali Salehi-Abari, and Julie Thorpe. A study on priming methods for graphical passwords. *Journal of Information Security and Applications*, 62:102913, 2021.

[32] Sarah Pearman, Shikun Aerin Zhang, Lujo Bauer, Nicolas Christin, and Lorrie Faith Cranor. Why people (don't) use password managers effectively. In *USENIX SOUPS*, 2019.

[33] Hirak Ray, Flynn Wolf, Ravi Kuber, and Adam J Aviv. Why older adults (don't) use password managers. In *USENIX SOUPS*, 2021.

[34] Elissa M Redmiles, Yasemin Acar, Sascha Fahl, and Michelle L Mazurek. A summary of survey methodology best practices for security and privacy researchers. Technical report, 2017.

[35] Elissa M. Redmiles, Sean Kross, and Michelle L. Mazurek. How well do my results generalize? comparing security and privacy survey results from mturk, web, and telephone samples. In *IEEE Symposium on Security and Privacy (SP)*, pages 1326–1343, 2019.

[36] Karen Renaud, Verena Zimmerman, Joseph Maguire, and Steve Draper. Lessons learned from evaluating eight password nudges in the wild. In *The LASER Workshop: Learning from Authoritative Security Experiment Results (LASER 2017)*, pages 25–37, 2017.

[37] Karen Renaud and Verena Zimmermann. Nudging folks towards stronger password choices: providing certainty is the key. *Behavioural Public Policy*, 3(2):228–258, 2018.

[38] Karen Renaud and Verena Zimmermann. Encouraging password manager use. *Network Security*, 2019(6):20–20, 2019.

[39] Sunyoung Seiler-Hwang, Patricia Arias-Cabarcos, Andres Marin, Florina Almenares, Daniel Diaz-Sanchez, and Christian Becker. I don't see why I would ever want to use it, analyzing the usability of popular smartphone password managers. In *ACM Conference on Computer and Communications Security (CCS)*, pages 1937–1953, 2019.

[40] James Simmons, Oumar Diallo, Sean Oesch, and Scott Ruoti. Systematization of password manager use cases and design paradigms. In *Annual Computer Security Applications Conference*, pages 528–540, 2021.

[41] Elizabeth Stobert and Robert Biddle. A password manager that doesn't remember passwords. In *New Security Paradigms Workshop*, pages 39–52, 2014.

[42] Elizabeth Stobert, Tina Safaie, Heather Molyneaux, Mohammad Mannan, and Amr Youssef. Bypass: Reconsidering the usability of password managers. In *International Conference on Security and Privacy in Communication Systems*, pages 446–466. Springer, 2020.

[43] Kurt Thomas, Jennifer Pullman, Kevin Yeo, Ananth Raghunathan, Patrick Gage Kelley, Luca Invernizzi, Borbala Benko, Tadek Pietraszek, Sarvar Patel, Dan Boneh, et al. Protecting accounts from credential stuffing with password breach alerting. In *28th USENIX Security Symposium*, pages 1556–1571, 2019.

[44] Julie Thorpe, Muath Al-Badawi, Brent MacRae, and Amirali Salehi-Abari. The presentation effect on graphical passwords. In *ACM Conference on Human Factors in Computing Systems (CHI)*, pages 2947–2950, 2014.

[45] Anthony Vance, David Eargle, Kirk Ouimet, and Detmar Straub. Enhancing password security through interactive fear appeals: A web-based field experiment. In *IEEE International Conference on System Sciences*, pages 2988–2997, 2013.

[46] Tina AG Venema, Floor M Kroese, Jeroen S Benjamins, and Denise TD De Ridder. When in doubt, follow the crowd? responsiveness to social proof nudges in the absence of clear preferences. *Frontiers in psychology*, 11:1385, 2020.

[47] Shengqian Wang, Amirali Salehi-Abari, and Julie Thorpe. Pixi: Password inspiration by exploring information. *arXiv preprint arXiv:2304.10728*, 2023.

[48] Yang Wang, Pedro Leon, Kevin Scott, Xiaoxuan Chen, Alessandro Acquisti, and Lorrie Cranor. Privacy nudges for social media: an exploratory facebook study. In *International Conference on World Wide Web*, 2013.

[49] Jeff Yan and Dearbhla McCabe. Gender bias in password managers. *arXiv e-prints*, pages arXiv–2206, 2022.

[50] Iryna Yevseyeva, Charles Morisset, and Aad van Moorsel. Modeling and analysis of influence power for information security decisions. *Performance Evaluation*, 98:36–51, 2016.

[51] Samira Zibaei, Dinah Rinoa Malapaya, Benjamin Mercier, Amirali Salehi-Abari, and Julie Thorpe. Do password managers nudge secure (random) passwords? In *USENIX SOUPS*, 2022.

[52] Verena Zimmermann and Karen Renaud. The nudge puzzle: matching nudge interventions to cybersecurity decisions. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 28(1):1–45, 2021.

Appendix A Demographic Information

Table 9: Demographic information across all conditions.

		Condition					
		Chrome	**Safari**	**PwdField**	**PwdField-No-Visual**	**Pop-up**	**Safari-Social**
		n=139	n=143	n=153	n=142	n=126	n=150
Gender	Female	28.1%	40.6%	45.1%	46.5%	38.1%	44.9%
	Male	71.2%	58.7%	54.2%	53.5%	61.1%	54.4%
	N/A	0.7%	0.7%	0.7%	0.0%	0.8%	0.7%
Age	18-25	19.4%	15.4%	20.9%	24.6%	24.7%	18.8%
	26-35	63.3%	49.6%	45.1%	42.3%	46.8%	34.1%
	36-50	14.4%	23.8%	22.8%	26.8%	21.4%	31.9%
	50+	2.2%	10.5%	10.5%	6.3%	7.1%	14.5%
	N/A	0.7%	0.7%	0.7%	0.0%	0.0%	0.7%
Education	High school	5.1%	9.1%	11.8%	4.2%	8.7%	10.9%
	Bachelor's	71.9%	67.8%	64.1%	75.4%	61.1%	70.3%
	Master's	19.4%	21.0%	22.1%	19.7%	30.2%	17.4%
	PhD/higher	2.2%	1.4%	1.3%	0.0%	0.0%	0.7%
	N/A	1.4%	0.7%	0.7%	0.7%	0.0%	0.7%
Study/Work	Social Sci.& Humanities	8.7%	4.2%	5.2%	8.5%	5.6%	10.2%
	Science	2.2%	5.6%	1.3%	0.7%	4.0%	5.3%
	Health Science	8.0%	16.1%	19.0%	12.0%	17.5%	12.4%
	Engineering & Applied Sci.	12.3%	7.0%	9.8%	7.7%	7.1%	4.3%
	Energy & Nuclear Sci.	1.4%	1.4%	2.0%	1.4%	1.6%	2.2%
	Education	9.4%	10.4%	6.5%	6.3%	7.8%	9.4%
	Business & IT	53.8%	49.7%	49.0%	57.0%	50.0%	47.8%
	Other	2.8%	3.5%	6.5%	6.4%	5.6%	7.0%
	N/A	1.4%	2.1%	0.7%	0.0%	0.8%	1.4%
Language	English	100.0%	97.2%	99.3%	100%	98.4%	97.1%
	Other	0.0%	2.1%	0.0%	0.0%	1.6%	2.2%
	N/A	0.0%	0.7%	0.7%	0.0%	0.0%	0.7%