Prospects for Improving Password Selection

Joram Amador, Yiran Ma, Summer Hasama, Eshaan Lumba,
Gloria Lee, and Eleanor Birrell, *Pomona College*

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

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.

Prospects for Improving Password Selection

Joram Amador
Pomona College

Yiran Ma
Pomona College

Summer Hasama
Pomona College

Eshaan Lumba
Pomona College

Gloria Lee
Pomona College

Eleanor Birrell
Pomona College

Abstract

User-chosen passwords remain essential to online security, and yet users continue to choose weak, insecure passwords. In this work, we investigate whether *prospect theory*, a behavioral model of how people evaluate risk, can provide insights into how users choose passwords and whether it can motivate new designs for password selection mechanisms that will nudge users to select stronger passwords. We run a pair of online user studies, and we find that an intervention guided by prospect theory—which leverages the reference-dependence effect by framing a choice of a weak password as a loss relative to choosing a stronger password—causes approximately 25% of users to improve the strength of their password (significantly more than alternative interventions) and improves the strength of passwords users select. We also evaluate the relation between feedback provided and password decisions and between users' mental models and password decisions. These results provide guidance for designing and implementing password selection interfaces that will significantly improve the strength of user-chosen passwords, thereby leveraging insights from prospect theory to improve the security of systems that use password-based authentication.

1 Introduction

User-chosen passwords remain a critical component of security. Many efforts have been made to nudge users towards choosing stronger passwords, including password rules [33] and password meters [20], but these efforts have met with only partial success. Password rules are ineffective at enforcing strong password choices [33, 69], many password meters are ineffective [13] especially for accounts users consider unimportant [20], and users continue to select and use weak passwords [45]. In this work, we investigate the extent to which insights from behavioral economics apply to users' password selection decisions and how those insights might be leveraged to enhance security by nudging users to select stronger passwords.

Prospect theory [32, 60–63] is an empirically-grounded behavioral model of how people make decisions in the presence of risk. Prospect theory has been applied to various different areas of economics; it has proven a successful model both for explaining observed behaviors [8, 12, 16, 28, 37, 38, 44, 54, 55] and for prescriptively nudging people towards higher-utility choices [24, 29, 39, 59].

Interactions between humans and systems that affect security and privacy can be framed as decisions in the presence of risk. For example, password selection requires users to evaluate the risk associated with each possible password they consider (how likely is it that their account will be compromised if they select that password and how bad will the consequences be if that occurs) and balance that risk against other competing factors (e.g., memorability and easy of typing, including on mobile devices) in order to decide which password to use. However, prior work has thus far explored the intersection between prospect theory and security and privacy only in limited specific domains, such as investment in security [53, 66], adoption of two-factor authentication [48], disclosure of personal information [3, 4, 27], cookie consent [42], and tracking authorization [18]. In this work, we explore the connection between prospect theory and password selection through a pair of online user studies.

Our first study explores the connection between two effects identified in the prospect theory literature—the *reference-dependence effect* and the *source-dependence effect*—and password selection. We ran an online user study on Amazon Mechanical Turk with 762 participants in which we asked people to create an account on an experimental website. Users who initially selected weak passwords or moderate passwords

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

were presented with an interactive prompt asking whether they wanted to go back and choose a stronger password; there were six different versions of the interactive prompt corresponding to three different framings (positive, neutral, and negative) and two different prompt phrasings (specific and vague). Participants also completed a follow-up survey about their beliefs regarding passwords and password-related risks. We found that the reference-dependence effect applies to password selection decisions—i.e., an interaction with negative framing resulted in significantly higher rates of improvement compared to neutral framing ($p < .001$) or positive framing ($p = .027$). However, the source-dependence effect did not appear to apply; the phrasing of the prompt (specific of vague) did not have a significant impact on whether user went back and selected a stronger password.

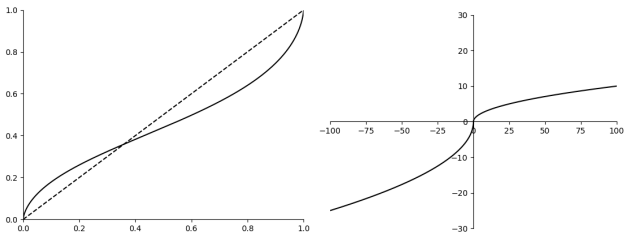
To validate the reference-dependence effect and to further understand how it influences password selection, we conducted a second user study through Prolific ($n = 607$) in which we recorded fine-grained measurements about password strength—as measured by the `zxcvbn` password meter's estimate of the number of guesses it would take to crack each password—along with information about how people modified their passwords. We also explored the impact of feedback and suggestions by including a condition with no meter and conditions in which the interactive prompt included suggestions for improving the password. Our results validated the reference-dependence effect for password selection decisions and provided insight into how people change their passwords in response to such interventions. We did not observe any significant differences due to feedback or suggestions.

Finally, we investigated whether mental models of security affected how users responded to our interactions. We found that perceptions about likely targets are correlated with password selection decisions but that decisions were consistent across different models of risks.

Our results suggest that some prospect theory effects can provide a model for understanding users' password selection decisions. In particular, we found that an intervention that leverages negative framing can significantly strengthen passwords. We believe that this insight from prospect theory can form the foundation for designing and implementing password selection mechanisms that enhance security by nudging users to select stronger passwords.

2 Background: Prospect Theory

Prospect theory [32,60–63]—first introduced in the 1970s as a critique of the then-dominant expected utility theory [23,67]—is a descriptive model of decision making in the presence of risk. Expected utility theory—which asserts that a principal faced with a choice between two options will evaluate the expected utility of each outcome and then select the option with the higher expected utility—does not accurately predict human behavior observed in many experimental settings.



(a) Decision weight function w (b) Subjective value function ν

Figure 1: Example functions matching empirically-observed behavior proposed by prior work [7, 63].

Prospect theory instead posits that decisions are comprised of two phases: an editing phase and an evaluation phase. In the editing phase, humans apply a set of simplifying heuristics to reduce the complexity of the decision problem. In the evaluation phase, probabilities and utilities are weighted by a decision weight w and a subjective value ν, respectively; example functions capturing empirically-observed behavior are shown in Figure 1. Humans are then presumed to rationally evaluate the options based on the weighted expected subjective value of the edited prospects.

The interactions between the editing phase and the weighting functions w and ν result in several effects that have been empirically validated through a series of experimental studies:

1. *Isolation Effect:* People simplify decision problems by disregarding components shared between alternatives and focusing exclusively on components that distinguish the options.

2. *Pseudocertainty Effect:* People simplify decision problems by treating extremely likely (but uncertain) outcomes as though they were certain.

3. *Reference-dependence Effect:* People simplify decision problems by defining outcomes relative to a neutral baseline. The framing of a problem can effect which baseline is used.

4. *Certainty Effect:* People overweight the probability of outcomes that are certain relative to outcomes that are merely probable.

5. *Source-dependence Effect:* People have different decision weights depending on the type of risk. For example, people have higher decision weights for contingent risks than for equivalent probabilistic risks (e.g., they prefer an insurance policy that provides certain coverage of specific types of damages to one that provides probabilistic coverage of all types of damages). Similarly, people are *ambiguity averse*—they prefer to bet based on precisely defined odds rather than on unknown odds.

6. *Loss Aversion Effect:* People subjectively dislike losses more than they value gains. That is, the value function is steeper for negative values (losses) than for positive values (gains).

More than 40 years later, prospect theory is still widely viewed as the best available model for how people make decisions in the presence of risk. It has been applied as a descriptive model to explain observed behavior in various different areas of economics including finance [6, 19, 44, 54], insurance [8, 30, 37, 56], savings [38], price setting [28], labor supply [12, 16], and betting markets [55]. Within the domain of computer science, prospect theory has been applied to explain decisions relating to investment in security [53, 66], adoption of two-factor authentication [48], disclosure of personal information [3, 4, 27], cookie consent [42], and tracking authorization [18].

Prospect theory has also been applied prescriptively in certain domains to nudge people towards certain "desirable" behaviors, including nudging employees to increase their retirement contributions [59], encouraging teachers to improve student outcomes [24], and incentivizing teams in high-tech factories to increase their productivity [39]. However, prospect-driven interventions have not been uniformly successful: a 2012 study did not see any increase in effort when financial or non-financial incentives for students were framed as losses compared to equivalent incentives framed as gains [29], and a 2021 study found that framing did not significant effect user decisions about whether to authorize tracking by iOS apps.

3 Related Work

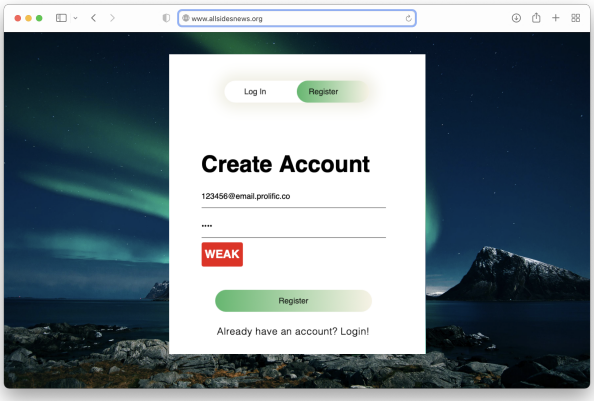
Improving Password Selections. Given the prevalence of password-based authentication and the ongoing dependence on user-chosen passwords, a large body of work has been dedicated to improving the strength of passwords that users select.

Early work on estimating password strength generally focused on entropy-based metrics [36]. However, entropy has since been criticized as been a poor measure of password guessability [33, 69, 70]. More recent efforts use dictionaries of words, lists of leaked passwords, and variants of words in those dictionaries and lists (e.g., L33t-style substitutions or addition of common suffixes) to define classes of weak or prohibited passwords [25, 43, 70].

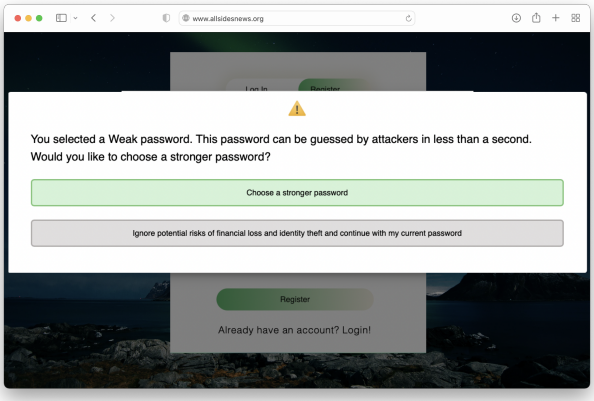
Studies have found that users exhibit misconceptions about password strength [65], which has resulted in increasing adoption of password meters across the most popular websites [20]. In general, having a password meter improves password strength, especially for accounts that users consider important [20]. However, some websites continue to use metrics that rely on entropy-based metrics and are thus inconsistent at effecting strong password selections [69]; one study found that most password meters deployed on actual websites are ineffective [13]. Careful calibration is also required to ensure that usability considerations do not undermine the benefits of a password meter: meters that are too strict can annoy users, while meters that are too lenient can result in weaker password selections [64].

Applications of Prospect Theory to Security. Despite the success of prospect theory in economics, there has been limited work applying prospect theory to security decisions, and only in limited domains. Verendel [66] developed a prospect theory model for decisions about buying versus skipping security protections (e.g., anti-virus software), although that work did not include any experimental validation. Schroeder [53] conducted a lab-based survey of IT officers in the U.S. military and found that prospect theory predicted hypothetical decisions about investment in information security. Sawicka and Gonzalez [51] explored the extent to which prospect theory can explain behavioral dynamics in IT-based work environments; they found the model matched choices observed in a short experimental run, but that it was not likely to account accurately for behavior over longer time periods. Sanjab et al. [50] explored how the decision weight function and value function impact principals' decisions in adversarial games in the context of attacks on Unmanned Aerial Vehicles (UAVs); they found that these subjective functions led to the adoption of riskier strategies, which cause delays in delivery. Most recently, Qu et al. [48] investigated the reference-dependence effect and the pseudocertainty effect in the context of two-factor authentication; they found that both effects explained whether or not users choose to enable two-factor authentication for a game in a laboratory setting. However, other security decisions—notably including password selection—have not been previously studied.

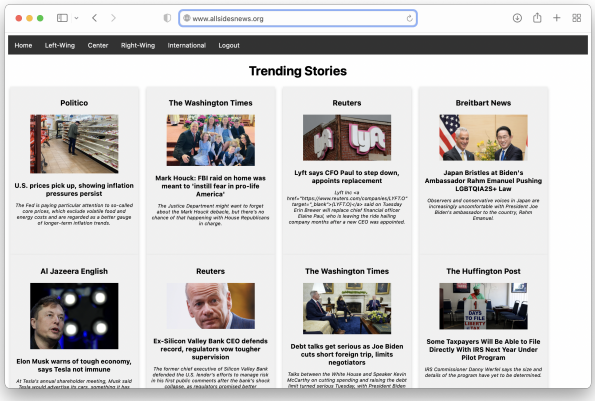
Applications of Prospect Theory to Privacy. In 2007, Acquisti et al. posited that several prospect theory effects—notably ambiguity aversion—might significantly impact privacy decision making [2]. Follow-up work found that people were more willing to sell personal information than to buy back previously-disclosed information [3, 27], and that the framing of notices affected whether or not users disclosed personal information in a survey [4]. Chloe et al. [14] also found that visual signals of an app's trustworthiness were affected by framing, but found that *positively* framed signals were more effective at nudging users away from low-privacy apps. More recent work has looked at developing and validating a theory for how context and personality affect decisions about disclosing personal information [5] and at the mechanism-design problem of how to calibrate noise in privacy-preserving mechanisms [40, 41].



(a) Account creation page



(b) Example interactive prompt



(c) Website home page

Figure 2: Screenshots of the account creation process on the example site

4 Methodology

To investigate how well prospect theory effects apply as a descriptive model of password selection, we conducted a pair of online user studies to evaluate the impact of two prospect theory effects—the source-dependence effect and the reference-dependence effect—on password selection decisions. These studies also explored the impact of feedback and mental models on password selection.

4.1 Experimental Setup

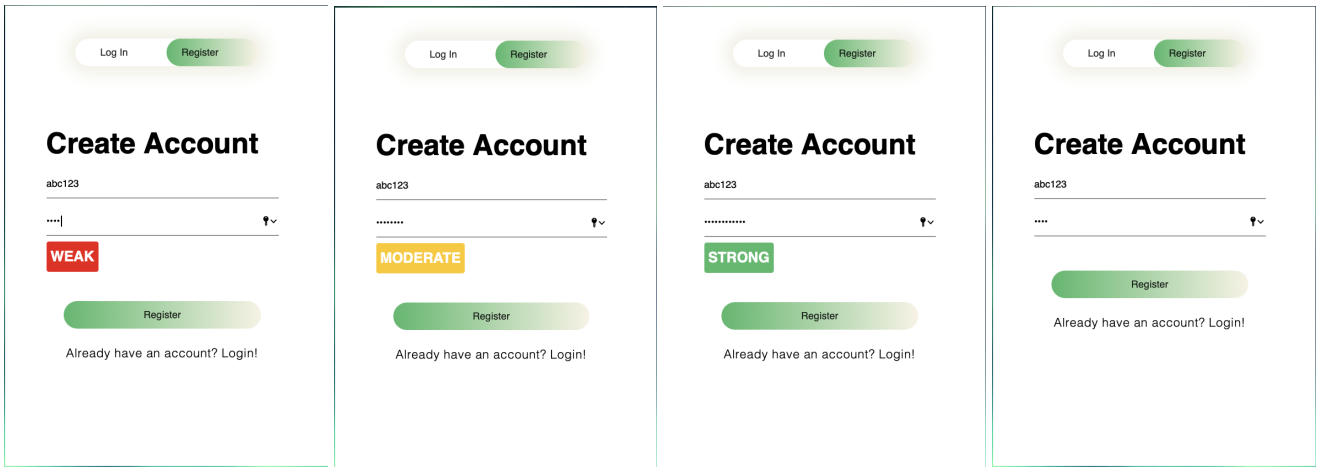
We developed an experimental aggregated news site that is accessible only to authenticated users. When visiting for the first time, each user is pseudorandomly assigned to a condition based on a hash of their current IP address. The user is then redirected to a condition-specific version of the account creation page (Figure 2a).

The initial account creation page had two different versions:

1. *Password Meter:* In these conditions, the password strength is classified in real time using the `zxcvbn` password strength estimator [70], and this information is displayed to the user by a password meter. Each password is classified as weak if it has a `zxcvbn` total score of 0 or 1 and moderate if the password has a total score of 2. Passwords with a total score of 3 or 4 are considered strong. Screenshots showing examples of how this meter looks with weak, moderate, and strong passwords are shown in Figures 3a, 3b, and 3c.

2. *No Meter:* In this condition, no information is displayed to the user about the strength of their password. This condition is shown in Figure 3d.

All participants in User Study 1 and most participants in User Study 2 saw an account creation page with a password meter. To provide a baseline for exploring the impact of feedback on password selection decisions, User Study 2 also included a condition with no meter.



(a) Meter (weak password) (b) Meter (moderate password) (c) Meter (strong password) (d) No Meter

Figure 3: Screenshots depicting the initial account creation page in different conditions with different strength passwords.

After initially selecting a password, users who select a strong password are redirected to the home page of the aggregated news site (Figure 2c). Users who select a weak or moderate password are instead presented with an interactive prompt that states that weak (resp., moderate) passwords put their account at risk and asks whether they would like to choose a stronger password (Figure 2b).

This prompt was presented using one of four possible wordings:

1. *Vague Prompt:* The password you selected is ⟨*strength*⟩. Would you like to choose a stronger password?

2. *Specific Prompt:* ⟨*strength*⟩ passwords can be guessed or learned by attackers in ⟨*time*⟩, which may lead to the loss of personal information, including credit card info, and identity theft. Would you like to choose a stronger password?

3. *Moderate Prompt:* The password you selected is ⟨*strength*⟩. This password can be guessed by attackers in ⟨*time*⟩. Would you like to choose a stronger password?

4. *Moderate Prompt + Suggestions:* The password you selected is ⟨*strength*⟩. This password can be guessed by attackers in ⟨*time*⟩. Would you like to choose a stronger password? Suggestions to improve password: ⟨*suggestions*⟩

Here ⟨*strength*⟩ is the classification based on the `zxcvbn` total score of the password the user submitted: "Weak" or "Moderate". (Recall that users who submit a strong password are authenticated immediately and are not presented with a prompt.) For the moderate and specific prompts, ⟨*time*⟩ is the `zxcvbn` estimate for how long it would take to crack the password with an offline guessing attack if passwords are hashed and salted using a slow hashing algorithm with a moderate work factor (e.g., bcrypt, scrypt, or PBKDF2). We used the human-readable text generated by `zxcvbn`, for example, "less than a second", "20 seconds", or "5 minutes". ⟨*suggestions*⟩ used the the natural-language text suggestions generated by `zxcvbn`, which provided password-specific suggestions such as "Avoid repeated words and characters", "Add another word or two. Uncommon words are better", and "Predictable substitutions like '@' instead of 'a' don't help very much". Participants in User Study 1 were shown either a vague prompt or a specific prompt. All participants in User Study 2 were shown a moderate prompt, with some conditions including suggestions and some not.

As shown in Figure 2b, the prompt has two buttons: one to go back (and choose a different password) and one to continue creating the account with the current password. This pair of buttons is labeled with one of three possible framings:

1. *Positive Framing:*
   Go Back: Choose a stronger password to reduce the risks of financial loss and identity theft
   Continue: Create account with current password

2. *Neutral Framing:*
   Go Back: Yes
   Continue: No

3. *Negative Framing:*
   Go Back: Choose a stronger password
   Continue: Ignore potential risks of financial loss and identity theft and create account with current password

The identified threats—financial loss and identity theft—were selected to maximize the appearance of risk within the password selection decision. However, we believe these risks are appropriate to this context. Prior work has established that password reuse attacks—in which attackers use leaked cre-

Study	Meter?	Prompt	Framing
1	Yes	Vague	Positive
1	Yes	Vague	Neutral
1	Yes	Vague	Negative
1	Yes	Specific	Positive
1	Yes	Specific	Neutral
1	Yes	Specific	Negative
2	No	Moderate	Neutral
2	Yes	Moderate	Positive
2	Yes	Moderate	Neutral
2	Yes	Moderate	Negative
2	Yes	Moderate+Suggestions	Positive
2	Yes	Moderate+Suggestions	Neutral
2	Yes	Moderate+Suggestions	Negative

Table 1: Conditions included in the two user studies.

dentials from low-value accounts such as news websites to attempt to access high-value accounts such as back accounts and emails—commonly occur online [9]. These attacks, which take advantage of the common practice of reusing credentials across multiple websites—are suspected behind several high-profile account compromises [31].

Each condition is defined by its meter setting (no meter vs. password meter), the wording of its interactive prompt (vague, specific, moderate, or moderate + suggestions), and its framing (postive, neutral, or negative). The conditions included in each of our two user studies are summarized in Table 1.

Users who elect to continue are redirected to the site homepage. Users who choose to go back stay on the account creation page until they select and submit a second password; they are then redirected to the site home page (no matter how strong their second selected password is). After spending a short time on the site, study participants returned to Qualtrics and completed a follow-up survey. Participants in the second user study were also asked to re-authenticate on the website after completing the survey. The full sets of questions for the follow-up surveys are provided in Appendix A and Appendix B.

The precise information recorded varied between the two studies. In the first user study, the experimental site logged the coarse-grained strength of each user's initial password choice (weak, moderate, or strong), how they interacted with the interactive prompt (if applicable), and the coarse-grained strength of their second password choice (if applicable). In the second user study, the site additionally logged the number of guesses it would take to crack each password (as estimated by `zxcvbn`), the length of each password, and (for users who selected a second password) the edit distance between the two passwords. In the second study, the site also stored the salted hash of the final password selected; this information was deleted after data collection was complete. Due to ethical and security concerns, we did not record any plaintext passwords.

4.2 Participant Recruitment

We recruited participants for both user studies online. All participants were presented with a consent form that informed them about what data would be collected and how that data would be used; only people who consented to these practices participated in a study. Our user studies, including all consent forms and survey instruments, were reviewed and approved in advance by the Pomona College Institutional Review Board.

User Study 1. For our first user study, participants were recruited through Amazon Mechanical Turk. Participation was restricted to United States residents who had completed at least 50 HITs with an approval rate of at least 95%.

The task was advertised as beta-testing an aggregated news site. Each participant was asked to (1) spend 1-2 minutes exploring the website as they would normally behave as an Internet user, (2) enter the unique confirmation code displayed when they visited the site, and (3) complete the follow-up survey questions. To avoid the appearance of collecting any personal information, users were given an email address to use during account creation.

Participants who did not enter a valid confirmation code, for whom we had no recorded log data, or who submitted irrelevant or incoherent responses to our free-response attention check question were excluded from the study. The 762 participants who completed the full study were compensated $1.20. Median completion time for this study was 5.15 minutes.

User Study 2. Due to increasing concerns about the external validity of studies conducted on Amazon Mechanical Turk [57], we elected to recruit participants for our second user study through Prolific. We recruited a gender-balanced, U.S. sample; following the methodology of Tang et al. [57], we did not further restrict participation.

Because Prolific is exclusively a platform for conducting studies, we advertised our second user study as a study about how people interact with websites instead of framing it as beta testing. Participants were given the same instructions as in User Study 1. However, since Prolific provides all users with an anonymous email tied to their Prolific ID—through which messages can be sent to the internal Prolific messaging system—we asked participants to use that email address to create their account.

We applied the same exclusion criteria in both studies. The 607 participants who successfully completed User Study 2 were compensated $2.50. The median completion time for the full task was 7.07 minutes. The higher compensation compared to User Study 1 was due to a longer estimated completion time combined an increase in California's minimum wage between the two studies.

The demographics of our study populations are summarized in Table 2.

Demographic		Study 1	Study 2	U.S.
Age	18-24	7.3	19.1	11.9
	25-34	33.6	35.3	17.9
	35-44	33.4	23.4	16.4
	45-59	19.8	16.5	24.4
	60-74	5.7	5.4	20.6
	75+	0.2	0.2	8.8
Race	White	76.1	77.8	74.4
	Black	10.4	11.4	13.9
	Asian	12.1	12.2	6.6
	Native Am.	3.1	2.4	1.5
	Other	2.3	2.3	5.2
Gender	Male	53.4	49.9	48.7
	Female	45.6	46.8	51.3
	N.B./other	1.0	3.3	-

Table 2: Study population demographics compared to the demographics of the United States, as published in the American Community Survey (ACS).

4.3 Study Limitations

In this study, our participants selected passwords for an experimental news site—one that they would likely not access or use beyond the scope of the study—rather than select passwords for a real-world account. This lack of realism might have affected password selection decisions. While prior work has shown that people select similar passwords in online studies compared to passwords selected for real accounts [21, 43], that work was conducted 10 years ago and focused exclusively on Amazon Mechanical Turk.

Moreover, this work looks specifically at the impact of rephrasing and re-framing risks incurred by password selection decisions. However, participants did not use personal email addresses in our studies; participants might therefore not have felt that their password selection decision put them at risk.

Finally, the particular threats emphasized in the experimental design—threats of financial loss and identity theft—might not have resonated with all participants, since the experimental website was an aggregated news site that did not direct collect and personal or financial information.

We discuss the validity of our results further in Section 6.

4.4 Hypotheses

To explore how prospect-driven interventions impact password selection, we identified and evaluated six hypotheses.

Source-dependence effect. When presented with the vague prompt, a user is required to evaluate options in the presence of multiple different sources of risk: in addition to reasoning about how likely it is that an attacker would target this site or this user, the user must evaluate uncertainties about how hard it would be for an attacker to guess their password and about what the potential consequence of password compromise might be. When presented with the specific prompt, some of these uncertainties—in particular how hard it would be for an attacker to guess their password and what attackers might do after they have learned a user's password—are eliminated in favor of more concrete risks.

The source-dependence effect observes that users evaluate different types of risk differently, and in particular that ambiguities are evaluated differently than more concrete risks. We therefore hypothesize that users will evaluate the the option to continue with their current (weak or moderate) password more negatively when presented with the specific prompt than with the vague prompt, resulting in stronger password selection after interacting with the specific prompt compared to the vague prompt.

Hypothesis 1: *Users' password selection decisions exhibit the* source-dependence effect*, that is users assigned to the specific prompt conditions are more likely to strengthen their password and will ultimately select stronger passwords compared to users assigned to the vague prompt conditions.*

Reference-dependence effect. In the conditions with positive framing, the option to go back is labeled as "Choose a stronger password to reduce the risks of financial loss and identity theft". By emphasizing the benefits of going back, this framing implicitly nudges the user to consider the option to continue as the neutral reference point and the option to go back as a choice with higher utility relative to that reference point. By contrast, the negative framing emphasizes the loss of utility ("potential risks of financial loss and identity theft") associated with continuing with the current password, thereby implicitly nudging the user to treat the option to go back as the neutral reference point and to evaluate continuing as a loss of utility relative to that reference point.

The reference-dependence effect implies that this difference in framing will cause users assigned to a positive framing condition to evaluate the difference between going back (i.e., choosing a stronger password) and continuing (i.e., submitting a weak password) as a positive *gain* in utility, whereas users assigned to a negative framing condition will evaluate the difference between continuing (i.e., submitting a weak password) and going back (i.e., choosing a stronger password) as a *loss* of utility. The loss aversion effect suggests that the subjective value function is steeper for (relative) losses than for (relative) gains. We therefore hypothesize that users will evaluate the option to continue with the current (weak) password more negatively in the negative framing conditions than the positive framing conditions—even though the two options have the same absolute utility in all conditions—resulting in stronger passwords selected after interacting with the negative framing prompt compared to the neutral and positive framing prompts.

Hypothesis 2: *Users' password selection decisions exhibit the* reference-dependence effect*, that is users assigned to a negative framing condition—which frames going back as the neutral baseline and continuing as a loss relative to that baseline—are more likely to strengthen their password and will ultimately select stronger passwords compared to users assigned to neutral or positive framing conditions.*

Feedback and Suggestions. Even after users decide to improve the strength of their password, the ability to successfully do so depends on knowing what constitutes a stronger password. The presence of a real-time, interactive password meter—which gives users course-grained feedback about the strength of their password—is one way to enable users to discover what might constitute a stronger password. Another approach would be to provide users with concrete suggestions for how they might strengthen their password.

Hypothesis 3: *Users who are shown a real-time password meter—which rates the current strength of their password—are more likely to strengthen their password and will ultimately select a stronger password compared to users who have no real-time information about their passwords strength.*

Hypothesis 4: *Users who are given concrete suggestions for how to improve their password are more likely to strengthen their password and will ultimately select a stronger password compared to users who are not shown suggestions.*

Mental Models of Hacking. Our user study concluded with a series of questions about participants' mental models of hacking and password security. One question we asked was who participants believe are the primary targets of password stealing attacks. Drawing on Wash's taxonomy of hacker mental models [68], we provided three possible answer: hackers target everyone equally, hackers primarily target rich people, and hackers primarily target users with special privileges (e.g., system administrators). We hypothesized that users who believe that hackers target everyone equally will consider themselves to be a more likely target compared to users with other mental models and will therefore be more sensitive to risks associated with password compromise.

Hypothesis 5: *Users who believe everyone is equally likely to be targeted by a password stealing attack will be more likely to strengthen their password and will ultimately select stronger passwords.*

We also asked participants questions designed to understand how they evaluated password-related risks. In particular, we asked how likely they believed a password attack would be to compromise their password if they selected a weak (resp., moderate, strong) password. We hypothesized that users' beliefs about risks associated with passwords would correlate with users' password selection decisions.

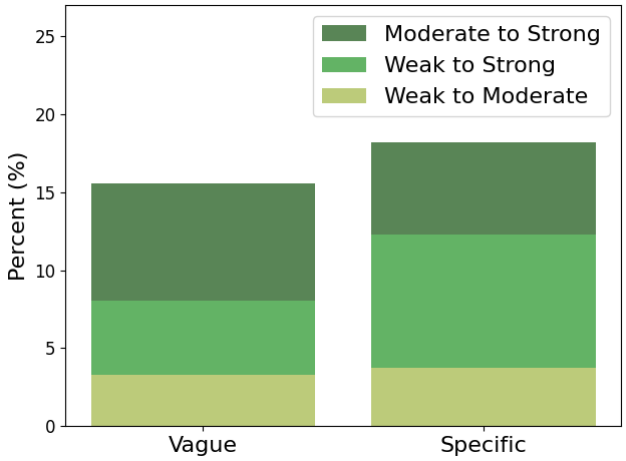


Figure 4: Percentage of users who improved password strength after interacting with vague and specific prompts.

Hypothesis 6: *Users who believe that weak passwords are more likely to be guessed by attackers will be more likely to initially choose a strong password, will be more likely to strengthen their password after seeing an interactive prompt, and will be more likely to ultimately choose a strong password.*

5 Results

To evaluate our hypotheses, we focused on users who initially selected weak or moderate passwords (i.e., users who saw the interactive prompt), and measured how many of those users (1) decided to go back and (2) selected a stronger password. We used χ^2-contingency tests to test for statistically significant differences.

5.1 Source-dependence Effect

To evaluate Hypothesis 1, we compared behavior in the conditions with vague wording to conditions with specific wording using data collected in User Study 1. We did not include conditions with moderate wording to avoid introducing confounding effects due to differences between study populations. We found that the specificity of the prompt had no significant effect on password selection. Of the users who saw a vague prompt, 15.6% opted to go back and ultimately selected a stronger password, compared to 18.2% of users who saw the specific prompt. This difference, depicted in Figure 4, was not statistically significant ($\chi^2 = .3$, $p = .573$).

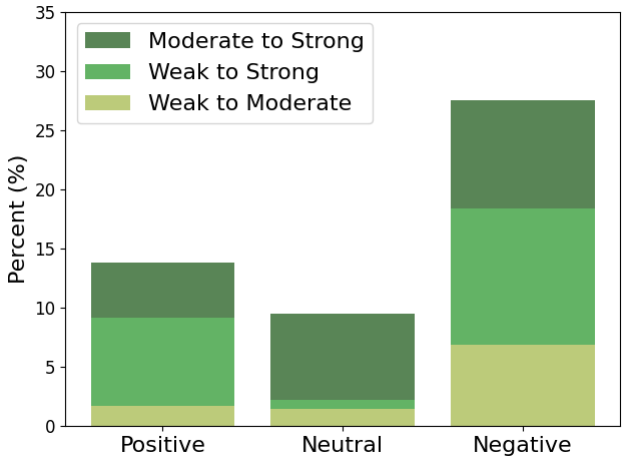


Figure 5: Percentage of users who improved password strength after seeing prompts with various framings.

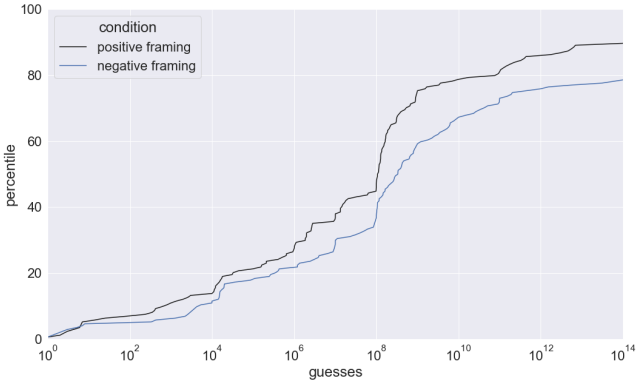


Figure 6: Strength of final password chosen by users who saw prompts with positive and negative framings, as measured by the percentile of final passwords in each condition that could be cracked with various numbers of guesses. Fewer passwords cracked corresponds to stronger passwords.

This negative result might be an indication that the source-dependence effect does not apply in the context of password selection decisions. However, it is also possible that the language of our prompts was insufficient to transfer uncertainty-based risk into probability-based risk in a manner that would trigger the source-dependence effect. Finally, it is possible that many of our users simply did not read the prompt, precluding the possibility of observing statistically significant effects due to the source-dependence effect.

Regardless of the underlying mechanism, these results suggest that utilizing more specific language about the nature of risks due to weak passwords—including notifying users of how long it would take an attacker to crack a password—is not an effective way to nudge users to select stronger passwords.

5.2 Reference-dependence Effect

To evaluate Hypothesis 2, we measured how many people strengthened their password after an intervention with negative framing compared to an intervention with positive framing (resp. neutral framing) using data collected in User Study 1. We conducted pairwise χ^2 tests to determine whether differences were significant. 25.9% of participants who saw a banner with negative framing went back and improved the strength of their password. This was significantly higher than the 14.2% who improved their password after seeing a banner with positive framing ($\chi^2 = 4.9$, $p = .027$) and the 9.5% who improved their password after seeing a banner with neutral framing ($\chi^2 = 11.5$, $p < .001$). There was no significant difference between the neutral framing and positive framing conditions ($\chi^2 = 1.0$, $p = .323$).

To validate that an interaction with negative framing improves password strength, we compared fine-grained strength—as measured by the estimated number of guesses it would take to crack that password—of the final password for our positive and negative framing conditions in User Study 2. We found that participants who saw a prompt with negative framing ultimately selected stronger passwords—i.e., passwords that would take more guesses for an attacker to crack—relative to participants who saw a prompt with positive framing (Figure 6).

These results suggest that the reference-dependence effect occurs in the context of password selection decisions. While further work will be required to validate this result in real-world systems, prior work has found that the results of password studies conducted online generally do extend to real-world systems [21]. The insight that the reference-dependence effect models users' password selection decisions therefore provides guidance for how authentication mechanisms designers might prescriptively enhance security: by adding a confirmation page and framing the option to go back and select a strong password as the "baseline" (and framing the option to continue with a weak or moderate password as a loss of utility relative to that baseline), we might be able to nudge users to enhance the security of their accounts by selecting significantly stronger passwords.

5.3 Feedback and Suggestions

To evaluate Hypothesis 3 and 4, we analyzed data from User Study 2 and compared conditions with a password meter and no suggestions to (1) our condition with no password meter and (2) our conditions that provided concrete suggestions for how to strengthen the initial password. In all cases, the fraction of participants who successfully strengthened their password was 22-26% (shown in Figure 7). There was no significant difference from removing the password meter ($\chi^2 < .1$, $p = .994$) or adding suggestions ($\chi^2 = .3$, $p = .606$).

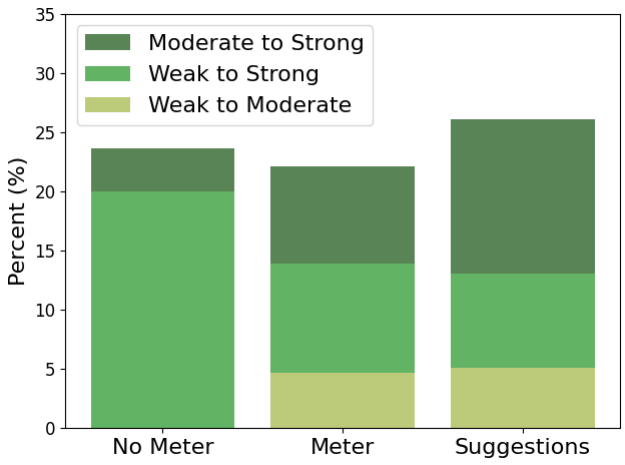


Figure 7: Percentage of users who improved passwords based on information provided about how to improve it.

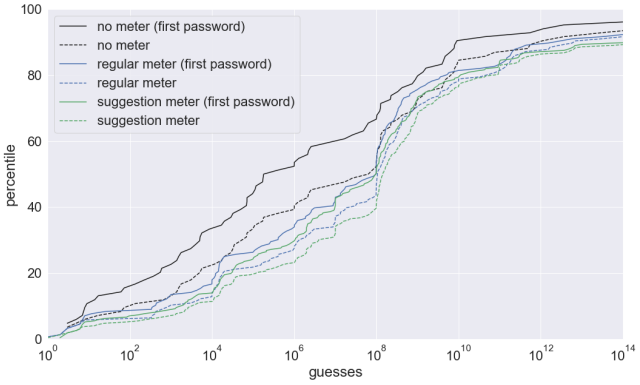


Figure 8: Impact of feedback and suggestions on password strength, measured by the percentile of passwords that could be cracked with various numbers of guesses. Fewer passwords cracked corresponds to stronger passwords.

To further understand this surprising negative result, we looked at data collected about fine-grained password strength—as defined by the number of guesses estimated to crack the password—both for the initial password selected and for the final password selected by participants in User Study 2. Like much prior work, we found that providing a password meter improved the strength of the initial password selected. Although final passwords selected in the no-meter condition were ultimately weaker than those selected in the conditions with a meter or a meter and suggestions, this distinction seems to be due to those weaker initial passwords rather than decreased ability to improve passwords without a password meter. We did not observe any significant differences between the conditions with and without suggestions, either in the strength of the initial password or in the strength of the ultimate password selected. These results are depicted in Figure 8.

Although we did not record plaintext passwords in either study, we did record password length and the edit distance between the two passwords (for people who selected a second password) for participants in User Study 2. We looked at this data to better understand the types of changes people made to their passwords. Overall, we found that 71% of people who decided to go back were successful at improving password strength. 8% stuck with their original password despite going back. 5% made small edits (defined as and edit distance of 3 or less) that did not improve strength. 13% made large edits that did not yield a stronger password; many of these large edits constituted selecting a completely new password. These results suggest that suggestions and feedback might be helpful for a minority of users; however, most people appear to already know how to strengthen their password and simply have to make the decision to do so. Definitively determining whether suggestions make prospect-driven interventions more effective will therefore require further work with larger sample sizes.

5.4 Mental Models

In our follow-up survey, we asked about users' mental models of password risks in order to explore whether there was a correlation between how users thought about password attacks and how users responded to our interactive prompts.

Hacking Targets. We asked participants in both user studies to identify who they thought hackers would target: everyone equally, primarily rich people, or primarily privileged users (e.g., system administrators). Overall, we found that 69.9% of participants believed that hackers target everyone equally and anyone is equally likely to have their password stolen, 10.6% of participants believed that hackers primarily target rich people, and 15.7% of participants believed that hackers primarily target privileged accounts (Figure 9).

A small number of participants opted instead to provide a free-form response. Some of these responses identified alternate groups as primary targets, including "gullible people", "weak links", and "older people". Other responses provided more nuanced variants of the options provided, e.g., "It depends on the hacker. Botnets attack everyone while social engineering attacks focus on special privileges" or identified all of the above as the best description of who is likely to be the target of a password attack.

Users who believed that hackers primarily target administrators during such attacks were significantly less likely to improve the strength of their password after exposure to the interactive prompt compared to users who believed that everyone is targeted equally ($\chi^2 = 4.9$, $p = .027$). We believe

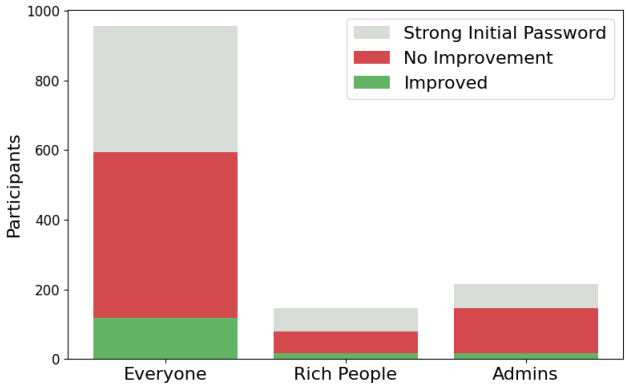


Figure 9: Password selection decisions broken down by perceived target of attacks.

this difference occurs because users with this mental model are less likely to believe they will be the target of an attack. To our surprise, users who believed that hackers target everyone equally were no more likely to improve the strength of their password compared to users who believed that hackers primarily target rich people ($\chi^2 = .1$, $p = .797$). This might be due to the fact that Americans consistently underestimate income inequality [17, 46, 47] and the income of top earners relative to the median worker [34] and thus might consider themselves to be a high-priority target even if they hold that mental model. Prior work has also found that participants recruited through Mechanical Turk and Prolific are more highly educated than the overall population [49, 57], a demographic that correlates with income and wealth.

Risk Evaluation. We also asked survey participants to rate how likely an attack would be to successfully compromise a password if a user selected (1) a weak password, (2) a moderate password, or (3) a strong password. We found that 88.1% of participants considered a weak password to be somewhat or very likely to be successfully attacked, compared to 53.7% of participants for a moderate password and 17.8% of participants for a strong password. These responses, which are depicted in Figure 10, were statistically significantly different between all the different password strengths ($\chi^2 \geq 382.4$, $p < .001$). These results suggest that most users believe that stronger passwords are in fact less likely to be vulnerable to password guessing attacks. However, there was no significant correlation between whether a user believed stronger passwords had less risk of being compromised and whether that user improved the strength of their password after seeing the interactive prompt.

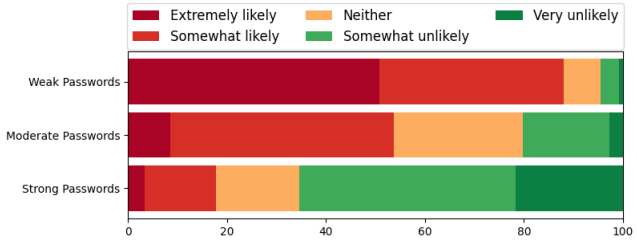


Figure 10: Perceptions of how likely a password guessing attack is to succeed based on the strength of a user's password.

6 Discussion and Limitations

Our results suggest that it might be possible to significantly improve the strength of user-chosen passwords by leveraging insights from prospect theory—in particular the reference-dependence effect—through a negatively-framed interactive prompt after users select an initial password. However, further work and careful consideration will be required to determine whether and how we should leverage these effects.

Ecological Validity. The major limitation of this work arises from the fact that we recruited participants through online crowdsourcing platforms to select passwords for an experimental account. Prior work has found that online study participants select slightly weaker passwords in experimental settings compared to real accounts. For example, one study found that 44.0% of users selected guessable passwords for their real account compared to 47.5% of Mechanical Turk users who were asked to select a password for an experimental study account given identical constraints [43]; in our first user study (also conducted on Mechanical Turk), we similarly found that 47.0% of our users initially selected a weak password (Figure 11). Despite these slight discrepancies, prior work has found that results from laboratory and online studies about passwords correspond to patterns in behavior for real accounts [21].

In addition to general threats to validity common to all online password studies, our focus on prospect theory—and the resulting need for participants to feel that their decisions might incur risk—introduces additional threats to validity. For ethical reasons, we did not collect any personal information (including personal email addresses) and instead had participants use dummy email addresses (User Study 1) or anonymous Prolific addresses (User Study 2), so participants might have realized that there was no actual risk incurred. This (lack of) realism might have influenced participants' decisions. Moreover, the choice of language in the warnings might have influenced how people reacted; some people might have disbelieved that weak passwords on news websites might incur financial risk. To explore how such confounding effects

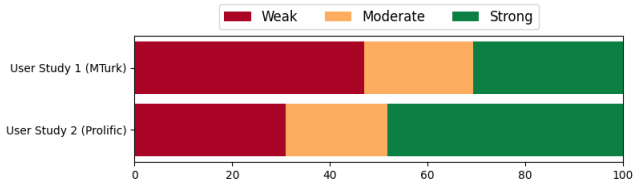


Figure 11: Strength of passwords initially selected by users.

might have affected our results, we asked participants in User Study 2 why they made the decision they made. About half of our participants responded in ways that suggested they were acting as though there were real risks (e.g., "Because I don't want to be at risk", "I didn't want my password to be too easy to guess", and "Its better to be safe than sorry"). However, other participants mentioned lack of realism (e.g., "This was not a real account and will not be using it again") or disbelief about the alleged risks (e.g., "Because I don't have any banking information on the website").

These results emphasize the importance of validating these effects (and their magnitude) in real-world contexts with actual risks. However, we hypothesize that the observed reference-dependent effect will extend to real-world password selection decisions, perhaps even with a larger effect size.

MTurk vs. Prolific. Recent work found that external validity of security and privacy surveys on Mechanical Turk has degraded over the last five years and that surveys conducted through Prolific now have higher external validity [57]. However, to our knowledge, this is the first work to conduct comparable experimental studies on both platforms.

Most results from our studies were consistent. People behaved similarly in both studies, e.g., significantly more people improved their password after seeing an intervention with negative framing (14.2% in User Study 1, 12.8% in User Study 2) than after seeing an intervention with negative framing (25.9% in User Study 1, 33.3% in User Study 2). Survey responses were also similar; e.g., 70.7% of participants in User Study 1 believed that attackers target everyone equally compared to 68.7% in User Study 2.

However, there was one notable difference between the studies: the strength of password people initially selected. Only 30.9% of participants in User Study 2 selected a weak password when presented with the same account creation interface used in User Study 1 compared to 47.0% in User Study 1 (Figure 11). This difference, which is statistically significant ($\chi^2 = 40.5$, $p < .001$) might be due to population differences between the two online platforms. Alternatively, it might be due to differences in how the study was presented. Since Prolific is a dedicated research study platform, participants in User Study 2 were aware all along that they were participating in a study rather than beta testing a website, which might have resulted in users selecting less realistic passwords. On the other hand, participants recruited through Prolific used a valid (albeit anonymous) email to create their account, perhaps resulting in participants selecting more realistic passwords. Differences between the observed rate of weak passwords in User Study 2 and that observed by prior work with real-world accounts might be symptomatic of low validity or might reflect temporal shifts in password selection behavior. Further research will be required to quantify the validity of experimental studies conducted on Prolific compared to Mechanical Turk and to validate password selection behavior in such studies today.

Memorability. The risk of account compromise due to password cracking and other attacks is not the only risk that users consider when selecting a password: users also need to weigh risks associated with other factors such as memorability. Forgetting a password is inconvenient in the best case; in the worst case, users can lose access to accounts. Future work will be required to determine the effect of framing on the memorability of passwords that users select.

Concerns about memorability might motivate users to employ memory-assistance techniques. This could lead to improved security practices—such as increased adoption of password managers—or to bad security practices—such as writing down passwords and leaving them in accessible locations. Further work will be required to evaluate the impact of framing on these other password-related practices.

Ethical Considerations. Leveraging the reference-dependence effect through negative framing of decisions has the potential to enhance security by encouraging users to adopt stronger passwords. However, this effect is an example of nudging [1]. While nudging is often associated with manipulative design elements that nudge users to make decisions that are inimical to their interests [10, 11, 15, 26], nudging can also be used towards making decisions that the mechanisms designer views as "better", a form of nudging sometimes called *soft paternalism* [22, 35, 52, 58]. Since nudging inherently leverages subconscious patterns in human behavior, care and consideration will be required to ensure that any prescriptive application of nudging and prospect theory effects with real-world impact—including leveraging the referenced-dependence effect to improve password selection—is handled ethically and responsibly.

Recommendations. Based on our results, we recommend adoption of password selection interfaces that present strong passwords as the default choice and follow-up prompts that emphasize risks associated with weak passwords. However, care will be required to ensure that this is ethically and effectively done without compromising memorability or other priorities. We recommend further research to to validate these results within real-world deployments and to ensure that the potential benefits to security outweigh any potential harms.

References

[1] Alessandro Acquisti, Idris Adjerid, Rebecca Balebako, Laura Brandimarte, Lorrie Faith Cranor, Saranga Komanduri, Pedro Giovanni Leon, Norman Sadeh, Florian Schaub, Manya Sleeper, Yang Wang, and Shomir Wilson. Nudges for privacy and security: Understanding and assisting users' choices online. *ACM Computing Surveys*, 50(3):1–41, 2017.

[2] Alessandro Acquisti, Stefanos Gritzalis, Costos Lambrinoudakis, and Sabrina di Vimercati. *What can behavioral economics teach us about privacy?* Auerbach Publications, 2007.

[3] Alessandro Acquisti, Leslie K. John, and George Loewenstein. What is privacy worth? *The Journal of Legal Studies*, 42(2):249–274, 2013.

[4] Idris Adjerid, Alessandro Acquisti, Laura Brandimarte, and George Loewenstein. Sleights of privacy: Framing, disclosures, and the limits of transparency. In *Symposium on Usable Privacy and Security*, pages 1–11, 2013.

[5] Gaurav Bansal, Fatemeh Mariam Zahedi, and David Gefen. Do context and personality matter? Trust and privacy concerns in disclosing private information online. *Information & Management*, 53(1):1–21, 2016.

[6] Nicholas Barberis and Ming Huang. Stocks as lotteries: The implications of probability weighting for security prices. *American Economic Review*, 98(5):2066–2100, 2008.

[7] Nicholas C. Barberis. Thirty years of prospect theory in economics: A review and assessment. *Journal of Economic Perspectives*, 27(1):173–96, 2013.

[8] Levon Barseghyan, Francesca Molinari, Ted O'Donoghue, and Joshua C Teitelbaum. The nature of risk preferences: Evidence from insurance choices. *American Economic Review*, 103(6):2499–2529, 2013.

[9] Bitglass. Where's your data? https://pages.bitglass.com/Bitglass_Where_is_your_Data_Report.html, 2016.

[10] Christoph Bösch, Benjamin Erb, Frank Kargl, Henning Kopp, and Stefan Pfattheicher. Tales from the dark side: Privacy dark strategies and privacy dark patterns. *Proc. Priv. Enhancing Technol.*, 2016(4):237–254, 2016.

[11] Harry Brignull and Alexander Darlo. Dark patterns. *Dark Patterns*, 2019.

[12] Colin Camerer, Linda Babcock, George Loewenstein, and Richard Thaler. Labor supply of New York City cabdrivers: One day at a time. *The Quarterly Journal of Economics*, 112(2):407–441, 1997.

[13] Xavier De Carné De Carnavalet and Mohammad Mannan. A large-scale evaluation of high-impact password strength meters. *ACM Transactions on Information and System Security (TISSEC)*, 18(1):1–32, 2015.

[14] Eun Kyoung Choe, Jaeyeon Jung, Bongshin Lee, and Kristie Fisher. Nudging people away from privacy-invasive mobile apps through visual framing. In *IFIP Conference on Human-Computer Interaction*, pages 74–91. Springer, 2013.

[15] Gregory Conti and Edward Sobiesk. Malicious interface design: Exploiting the user. In *Proceedings of the 19th International Conference on World Wide Web*, pages 271–280, 2010.

[16] Vincent P. Crawford and Juanjuan Meng. New York City cab drivers' labor supply revisited: Reference-dependent preferences with rational-expectations targets for hours and income. *American Economic Review*, 101(5):1912–32, 2011.

[17] Shai Davidai. Why do Americans believe in economic mobility? Economic inequality, external attributions of wealth and poverty, and the belief in economic mobility. *Journal of Experimental Social Psychology*, 79:138–148, 2018.

[18] Anzo DeGiulio, Hanoom Lee, and Eleanor Birrell. "Ask app not to track": The effect of opt-in tracking authorization on mobile privacy. In *Emerging Technologies for Authorization and Authentication (ETAA)*, pages 152–167. Springer, 2021.

[19] Stephen G. Dimmock and Roy Kouwenberg. Loss-aversion and household portfolio choice. *Journal of Empirical Finance*, 17(3):441–459, 2010.

[20] 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 *SIGCHI Conference on Human Factors in Computing Systems (CHI)*, pages 2379–2388, 2013.

[21] Sascha Fahl, Marian Harbach, Yasemin Acar, and Matthew Smith. On the ecological validity of a password study. In *Proceedings of the Ninth Symposium on Usable Privacy and Security*, pages 1–13, 2013.

[22] Bijan Fateh-Moghadam and Thomas Gutmann. Governing [through] autonomy. The moral and legal limits of "soft paternalism". *Ethical Theory and Moral Practice*, 17(3):383–397, 2014.

[23] Milton Friedman and Leonard J. Savage. The utility analysis of choices involving risk. *Journal of political Economy*, 56(4):279–304, 1948.

[24] Roland G. Fryer Jr, Steven D. Levitt, John List, and Sally Sadoff. Enhancing the efficacy of teacher incentives through loss aversion: A field experiment. Technical report, National Bureau of Economic Research, 2012.

[25] Maximilian Golla and Markus Dürmuth. On the accuracy of password strength meters. In *Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security*, pages 1567–1582, 2018.

[26] Colin M. Gray, Yubo Kou, Bryan Battles, Joseph Hoggatt, and Austin L. Toombs. The dark (patterns) side of UX design. In *Proceedings of the 2018 SIGCHI Conference on Human Factors in Computing Systems (CHI)*, pages 1–14, 2018.

[27] Jens Grossklags and Alessandro Acquisti. When 25 cents is too much: An experiment on willingness-to-sell and willingness-to-protect personal information. In *Workshop on Economics of Information Security*, 2007.

[28] Paul Heidhues and Botond Kőszegi. Regular prices and sales. *Theoretical Economics*, 9(1):217–251, 2014.

[29] Tanjim Hossain and John A. List. The behavioralist visits the factory: Increasing productivity using simple framing manipulations. *Management Science*, 58(12):2151–2167, 2012.

[30] Wei-Yin Hu and Jason S. Scott. Behavioral obstacles in the annuity market. *Financial Analysts Journal*, 63(6):71–82, 2007.

[31] 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.

[32] Daniel Kahneman and Amos Tversky. Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2):263–292, 1979.

[33] Patrick Gage Kelley, Saranga Komanduri, Michelle L Mazurek, Richard Shay, Timothy Vidas, Lujo Bauer, Nicolas Christin, Lorrie Faith Cranor, and Julio Lopez. Guess again (and again and again): Measuring password strength by simulating password-cracking algorithms. In *IEEE Symposium on Security and Privacy*, pages 523–537, 2012.

[34] Sorapop Kiatpongsan and Michael I. Norton. How much (more) should CEOs make? A universal desire for more equal pay. *Perspectives on Psychological Science*, 9(6):587–593, 2014.

[35] Gebhard Kirchgässner. Soft paternalism, merit goods, and normative individualism. *European Journal of Law and Economics*, 43(1):125–152, 2017.

[36] Saranga Komanduri, Richard Shay, Patrick Gage Kelley, Michelle L Mazurek, Lujo Bauer, Nicolas Christin, Lorrie Faith Cranor, and Serge Egelman. Of passwords and people: Measuring the effect of password-composition policies. In *SIGCHI Conference on Human Factors in Computing Systems*, pages 2595–2604, 2011.

[37] Botond Kőszegi and Matthew Rabin. Reference-dependent risk attitudes. *American Economic Review*, 97(4):1047–1073, 2007.

[38] Botond Kőszegi and Matthew Rabin. Reference-dependent consumption plans. *American Economic Review*, 99(3):909–36, 2009.

[39] Steven D. Levitt, John A. List, Susanne Neckermann, and Sally Sadoff. The behavioralist goes to school: Leveraging behavioral economics to improve educational performance. *American Economic Journal: Economic Policy*, 8(4):183–219, 2016.

[40] Guocheng Liao, Xu Chen, and Jianwei Huang. Optimal privacy-preserving data collection: A prospect theory perspective. In *IEEE Global Communications Conference*, pages 1–6, 2017.

[41] Guocheng Liao, Xu Chen, and Jianwei Huang. Prospect theoretic analysis of privacy-preserving mechanism. *Transactions on Networking*, 28(1):71–83, 2019.

[42] Yiran Ma and Eleanor Birrell. Prospective consent: The effect of framing on cookie consent decisions. In *SIGCHI Conference on Human Factors in Computing Systems Extended Abstracts*, pages 1–6, 2022.

[43] Michelle L Mazurek, Saranga Komanduri, Timothy Vidas, Lujo Bauer, Nicolas Christin, Lorrie Faith Cranor, Patrick Gage Kelley, Richard Shay, and Blase Ur. Measuring password guessability for an entire university. In *ACM Conference on Computer and Communications Security*, pages 173–186, 2013.

[44] Juanjuan Meng and Xi Weng. Can prospect theory explain the disposition effect? A new perspective on reference points. *Management Science*, 64(7):3331–3351, 2018.

[45] NordPass. Top 200 most common passwords of the year 2020. https://nordpass.com/most-common-passwords-list/.

[46] Michael I. Norton and Dan Ariely. Building a better america—one wealth quintile at a time. *Perspectives on psychological science*, 6(1):9–12, 2011.

[47] Michael I. Norton, David T. Neal, Cassandra L. Govan, Dan Ariely, and Elise Holland. The not-so-common-wealth of Australia: Evidence for a cross-cultural desire

for a more equal distribution of wealth. *Analyses of Social Issues and Public Policy*, 2014.

[48] Leilei Qu, Cheng Wang, Ruojin Xiao, Jianwei Hou, Wenchang Shi, and Bin Liang. Towards better security decisions: Applying prospect theory to cybersecurity. In *SIGCHI Conference on Human Factors in Computing Systems Extended Abstracts*, pages 1–6, 2019.

[49] 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*, pages 1326–1343, 2019.

[50] Anibal Sanjab, Walid Saad, and Tamer Başar. Prospect theory for enhanced cyber-physical security of drone delivery systems: A network interdiction game. In *IEEE International Conference on Communications*, pages 1–6, 2017.

[51] Agata Sawicka and Jose J. Gonzalez. Choice under risk in IT-environments according to cumulative prospect theory. In *21st International Conference of the System Dynamics Society, New York*, 2003.

[52] Jan Schnellenbach. Nudges and norms: On the political economy of soft paternalism. *European Journal of Political Economy*, 28(2):266–277, 2012.

[53] Neil J. Schroeder. Using prospect theory to investigate decision-making bias within an information security context. Technical report, Air Force Institution of Technology Wright-Patterson School of Engineering and Management, 2005.

[54] Hersh Shefrin and Meir Statman. The disposition to sell winners too early and ride losers too long: Theory and evidence. *The Journal of Finance*, 40(3):777–790, 1985.

[55] Erik Snowberg and Justin Wolfers. Explaining the favorite–long shot bias: Is it risk-love or misperceptions? *Journal of Political Economy*, 118(4):723–746, 2010.

[56] Justin Sydnor. (Over) insuring modest risks. *American Economic Journal: Applied Economics*, 2(4):177–99, 2010.

[57] Jenny Tang, Eleanor Birrell, and Ada Lerner. Replication: How well do my results generalize now? The external validity of online privacy and security surveys. In *Symposium on Usable Privacy and Security*, pages 367–385, 2022.

[58] Richard Thaler and Cass Sunstein. *Nudge: Improving Decisions About Health, Wealth, and Happiness*. Yale University Press, 2008.

[59] Richard H Thaler and Shlomo Benartzi. Save more tomorrow: Using behavioral economics to increase employee saving. *Journal of Political Economy*, 112(S1):S164–S187, 2004.

[60] Amos Tversky and Daniel Kahneman. The framing of decisions and the psychology of choice. *Science*, 211(4481):453–458, 1981.

[61] Amos Tversky and Daniel Kahneman. The framing of decisions and the evaluation of prospects. In *Studies in Logic and the Foundations of Mathematics*, volume 114, pages 503–520. Elsevier, 1986.

[62] Amos Tversky and Daniel Kahneman. Loss aversion in riskless choice: A reference-dependent model. *The Quarterly Journal of Economics*, 106(4):1039–1061, 1991.

[63] Amos Tversky and Daniel Kahneman. Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, 5(4):297–323, 1992.

[64] Blase Ur, Patrick Gage Kelley, Saranga Komanduri, Joel Lee, Michael Maass, Michelle L. Mazurek, Timothy Passaro, Richard Shay, Timothy Vidas, Lujo Bauer, Nicolas Christin, and Lorrie Faith Cranor. How does your password measure up? The effect of strength meters on password creation. In *21st USENIX Security Symposium*, pages 65–80, 2012.

[65] Blase Ur, Fumiko Noma, Jonathan Bees, Sean M Segreti, Richard Shay, Lujo Bauer, Nicolas Christin, and Lorrie Faith Cranor. "I added '!'at the end to make it secure": Observing password creation in the lab. In *Eleventh Symposium On Usable Privacy and Security*, pages 123–140, 2015.

[66] Vilhelm Verendel. *A prospect theory approach to security*. Citeseer, 2008.

[67] John von Neumann and Oskar Morgenstern. *Theory of Games and Economic Behavior*. Princeton University Press, 1944.

[68] Rick Wash. Folk models of home computer security. In *Proceedings of the Sixth Symposium on Usable Privacy and Security*, pages 1–16, 2010.

[69] Matt Weir, Sudhir Aggarwal, Michael Collins, and Henry Stern. Testing metrics for password creation policies by attacking large sets of revealed passwords. In *ACM Conference on Computer and Communications Security*, pages 162–175, 2010.

[70] Daniel Lowe Wheeler. zxcvbn: Low-budget password strength estimation. In *25th USENIX Security Symposium*, pages 157–173, 2016.

A Follow-up Survey Questions: User Study 1

1. Provide a brief description of the website you visited. (A few words or 1 sentence is sufficient.)

   [free response]

2. How strong was the password you chose when you created your account on the site?

   - Strong
   - Moderate
   - Weak

3. How much do you agree with the statement: A hacker would be likely to try to hack this site.

   - Completely agree
   - Somewhat agree
   - Neither agree nor disagree
   - Somewhat disagree
   - Completely disagree

4. How much do you agree with the statement: A hacker would be likely to successfully guess the password I used on this site.

   - Completely agree
   - Somewhat agree
   - Neither agree nor disagree
   - Somewhat disagree
   - Completely disagree

5. Is the password you used on this site a password that you also use on other sites?

   - Yes
   - No

6. How common are password stealing attacks?

   - Extremely common
   - Somewhat common
   - Neither common nor uncommon
   - Somewhat uncommon
   - Extremely uncommon

7. How could hackers potentially learn your password? Choose all that apply.

   - It is impossible for a hacker to learn my password.
   - If I accidentally download a virus, a malicious app, or a malicious attachment.
   - If I visit a sketchy or malicious website.
   - If I accidentally click on a phishing link and enter my credentials on a fake website.
   - If a hacker (or a program run by a hacker) guesses my password on the website.
   - If a hacker steals the files storing all passwords for the website.
   - Other: _____

8. How likely would it be for a password stealing attack to succeed if you use a weak password?

   - Extremely likely
   - Somewhat likely
   - Neither likely nor unlikely
   - Somewhat unlikely
   - Extremely unlikely

9. How likely would it be for a password stealing attack to succeed if you use a moderate password?

   - Extremely likely
   - Somewhat likely
   - Neither likely nor unlikely
   - Somewhat unlikely
   - Extremely unlikely

10. How likely would it be for a password stealing attack to succeed if you use a strong password?

    - Extremely likely
    - Somewhat likely
    - Neither likely nor unlikely
    - Somewhat unlikely
    - Extremely unlikely

11. Do you think upgrading your passwords can prevent password guessing?

    - Yes
    - Maybe
    - No

12. What could a hacker do if they successfully learn your password? Choose all that apply.

    - They could cause bugs (viruses can cause computers to crash, quit applications, erase important system files).
    - They could steal personal and financial information from individual computers, and send the information to criminal.
    - They could resell personal information.

- They could display annoying visual images on computers (a skull, advertising popups, or pornography).
- They could control the computer and use the computer to send information to others.
- They could use the computers to cause problems for third parties.
- Other: _____

13. Which of the following are likely to try to steal passwords? Choose all that apply.
    - A young computer geek who wants to show off or explore the internet
    - Criminals
    - Organizations and institutions
    - Other: _____
14. Which of the following best describes who is likely to be the target of a password stealing attack?
    - Hackers target everyone equally, and anyone is equally likely to have their password stolen
    - Hackers primarily target rich people
    - Hackers primarily target people with special privileges (e.g, system administrators)
    - Other: _____
15. What is your current age?
    - 18-24
    - 25-34
    - 35-44
    - 45-59
    - 60-74
    - 75+
16. What is your gender?
    - Man
    - Woman
    - Non-binary person
    - Other: _____
17. Choose one or more races that you consider yourself to be:
    - White
    - Black or African American
    - American Indian or Alaska Native
    - Asian
    - Pacific Islander or Native Hawaiian
    - Other: _____
18. Do you consider yourself to be Hispanic?
    - Yes
    - No

B Follow-up Survey Questions: User Study 2

1. Provide a brief description of the website you visited. (A few words or 1 sentence is sufficient.)

   [free response]

2. After you initially entered your Prolific ID and password, did you see a notice that looked like this (see picture above)?
    - Yes
    - No
    - I don't remember

   [Questions 3-8 were displayed only if participant answered yes to Question 2.]

3. In your opinion, how high would the risk of account compromise, financial loss, or identity theft be if you continued with your initial password?
    - Very high risk
    - High risk
    - Moderate risk
    - Low risk
    - Very low risk
4. In your opinion, how high would the risk of forgetting your password or losing access to the account be if you continued with your initial password?
    - Very high risk
    - High risk
    - Moderate risk
    - Low risk
    - Very low risk
5. In your opinion, how high would the risk of account compromise, financial loss, or identity theft be if you went back and chose a stronger password?
    - Very high risk
    - High risk
    - Moderate risk

- Low risk
- Very low risk

6. In your opinion, how high would the risk of forgetting your password or losing access to the account be if you went back and chose a stronger password?
   - Very high risk
   - High risk
   - Moderate risk
   - Low risk
   - Very low risk
7. Which choice was presented as the default option?
   - [Option to choose a stronger password. Exact wording depended on condition.]
   - [Option to continue with current password. Exact wording depended on condition.]
   - I don't remember.
8. Why did you choose that option?

   [free response]

   [Questions 9-16 were displayed only if participant answered selected the option to choose a stronger password for Question 7.]
9. In your opinion, how strong was the first password you chose?
   - Strong
   - Moderate
   - Weak
10. In your opinion, how likely is it that a hacker would be able to guess the first password you chose?
    - Very likely
    - Somewhat likely
    - Neither likely nor unlikely
    - Somewhat unlikely
    - Very unlikely
11. In your opinion, how likely is it that you would forget the first password you chose?
    - Very likely
    - Somewhat likely
    - Neither likely nor unlikely
    - Somewhat unlikely
    - Very unlikely
12. Is the first password you chose one that you use on another website or account?
    - Yes
    - No
    - No, but I use similar passwords for other websites or accounts
    - Prefer not to say
13. In your opinion, how strong was the second password you chose?
    - Strong
    - Moderate
    - Weak
14. In your opinion, how likely is it that a hacker would be able to guess the second password you chose?
    - Very likely
    - Somewhat likely
    - Neither likely nor unlikely
    - Somewhat unlikely
    - Very unlikely
15. In your opinion, how likely is it that you would forget the second password you chose?
    - Very likely
    - Somewhat likely
    - Neither likely nor unlikely
    - Somewhat unlikely
    - Very unlikely
16. Is the second password you chose one that you use on another website or account?
    - Yes
    - No
    - No, but I use similar passwords for other websites or accounts
    - Prefer not to say

    [Questions 17-20 were displayed if participant answered No to Question 2 or selected the option to choose a stronger password for Question 7.]
17. In your opinion, how strong was the password you chose?
    - Strong
    - Moderate
    - Weak

18. In your opinion, how likely is it that a hacker would be able to guess the password you chose?
    - Very likely
    - Somewhat likely
    - Neither likely nor unlikely
    - Somewhat unlikely
    - Very unlikely
19. In your opinion, how likely is it that you would forget the password you chose?
    - Very likely
    - Somewhat likely
    - Neither likely nor unlikely
    - Somewhat unlikely
    - Very unlikely
20. Is the password you chose one that you use on another website or account?
    - Yes
    - No
    - No, but I use similar passwords for other websites or accounts
    - Prefer not to say
21. How much do you agree with the statement: Having strong passwords is important to me.
    - Strongly disagree
    - Disagree
    - Neutral
    - Agree
    - Strongly agree
22. How much do you agree with the statement: Having passwords I will remember is important to me.
    - Strongly disagree
    - Disagree
    - Neutral
    - Agree
    - Strongly agree
23. In general, how likely is it that a hacker would be able to guess a strong password?
    - Very likely
    - Somewhat likely
    - Neither likely nor unlikely
    - Somewhat unlikely
    - very unlikely
24. In general, how likely is it that a hacker would be able to guess a moderate password?
    - Very likely
    - Somewhat likely
    - Neither likely nor unlikely
    - Somewhat unlikely
    - Very unlikely
25. In general, how likely is it that a hacker would be able to guess a weak password?
    - Very likely
    - Somewhat likely
    - Neither likely nor unlikely
    - Somewhat unlikely
    - Very unlikely
26. Which of the following best describes whose passwords a hacker would try to learn?
    - Hackers target everyone equally, and anyone is equally likely to have their password stolen
    - Hackers primarily target rich people
    - Hackers primarily target people with special privileges (e.g, system administrators)
    - Other: _____
27. What could a hacker potentially do if they successfully learn your password? Choose all that apply.
    - They could cause bugs (viruses can cause computers to crash, quit applications, erase important system files).
    - They could steal personal and financial information from individual computers, and send the information to criminal.
    - They could resell personal information.
    - They could display annoying visual images on computers (a skull, advertising popups, or pornography).
    - They could control the computer and use the computer to send information to others.
    - They could use the computers to cause problems for third parties.
    - Other: _____

[Participants were then asked to return to the website and log-in again.]

28. Were you able to successfully log-in to your account on the website?
    - Yes, I remembered my password
    - Yes, but it took me multiple tries to remember my password
    - No, I was unable to log-in to my account

    [Question 29 was only displayed if the participant said they were able to log-in to their account in Question 28.]

29. How did you remember your password for this account?
    - I wrote it down
    - I used a password manager
    - I have used this password before
    - I just remembered it
    - Other: _____

30. How much do you agree with the statement: I am proficient with the Internet and computers?
    - Strongly disagree
    - Disagree
    - Neutral
    - Agree
    - Strongly agree

31. How much do you agree with the statement: I am knowledgeable about security and privacy?
    - Strongly disagree
    - Disagree
    - Neutral
    - Agree
    - Strongly agree

32. What is your current age?
    - 18-24
    - 25-34
    - 35-44
    - 45-59
    - 60-74
    - 75+

33. What is your gender?
    - Man
    - Woman
    - Non-binary person
    - Prefer not to say
    - Prefer to self-describe: _____

34. Choose one or more races that you consider yourself to be:
    - White
    - Black or African American
    - American Indian or Alaska Native
    - Asian
    - Pacific Islander or Native Hawaiian
    - Other: _____

35. Do you consider yourself to be Hispanic/Latinx/Latine?
    - Yes
    - No