The Decoy Effect for Passwords – A First Exploration

Tobias Seitz
Media Informatics Group
LMU Munich
Email: tobias.seitz@ifi.lmu.de

***Abstract*—This paper reports on an online survey eliciting effects of a decoy choice architecture for passwords. The survey measured preference for a given set of passwords and was conducted as a pre-study to a more interactive online-study to give us a feeling about the constraints of decoy passwords. The survey revealed that in some scenarios it is possible to influence the preference for a target password by introducing an unfavorable decoy password. However, the preference for our target password was generally high, which leads us to believe that the choice architecture needs to be evolved further.**

I. Introduction

Until passwords on the web are replaced by a more sophisticated authentication method, users face the challenge to manage many credentials. A common understanding among users is that their passwords need to be complex to be strong. Ur et al. showed that, consequently, users do not necessarily recognize a "strong" password when they see one [1]. We argue that providing feed*forward* instead of feedback can help users to identify stronger passwords than their own and potentially persuade them to reflect on their behavior.

In this paper, we report on our exploration of the decoy effect for passwords. As described in detail in Section II-A, the effect is often used in marketing to make a favorable option stand out against a competitor and a decoy option. We conducted an online survey to examine the phenomenon. We found that the decoy option was effective in some, but not in all of the voting tasks.

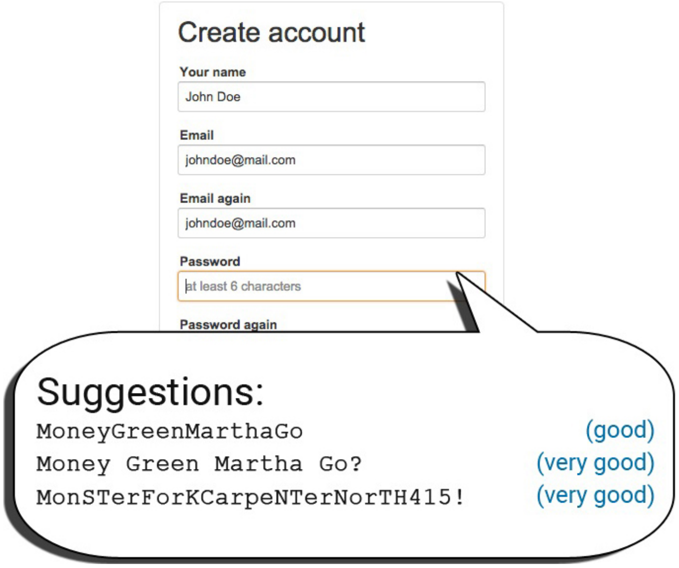
II. Background and Related Work

In this section we take a look at the particularities of the decoy effect and provide specific usage examples. Moreover, we discuss the feasibility of passphrases, which supposedly allowed us to elicit the preferences predicted by the decoy effect.

A. The Decoy Effect

When people face a choice between two items that can be ranked on two distinct dimensions, e.g. price and quality, adding another item can influence preference towards one of the original two items. This effect is found in consumer research and is often called the decoy effect. The three items are usually referred to as **competitor**, **target** and **decoy**. The competitor is commonly the least expensive option, which however is characterized by its low quality. Vendors are trying to boost sales for the target, because it has a greater margin and is of better quality than the competitor. The final decoy item can be constructed in many ways. For instance, it could be "just as good" as the target, but much more expensive (cf. Figure 4). In that case, buyers would not benefit at all from choosing it, thus taking the option would be irrational. In other scenarios the quality of the decoy might even be superior to the target's but it would not justify the increase in price for many consumers, again boosting the sales of the target (cf. Fig. 2). To illustrate the pricing and the framing of options, we take a closer look at one of the original examples by Huber et al. [2]. In some cases, participants in their studies were offered two options for a sixpack of beer:

Option	Price	Quality rating (0=worst, 100=best)
(A)	$1.50	50
(B)	$2.50	70



Create account

Your name
John Doe

Email
johndoe@mail.com

Email again
johndoe@mail.com

Password
at least 6 characters

Password again

Suggestions:

```
MoneyGreenMarthaGo                (good)
Money Green Martha Go?            (very good)
MonSTerForKCarpeNTerNorTH415!     (very good)
```

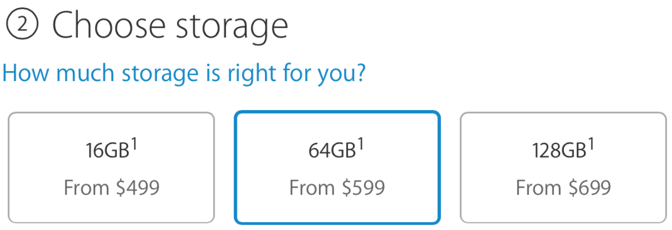
Fig. 1. We suggest three passphrases. The first one is the "competitor" and the weakest of all. The second is the "target"; it is readable, reasonably long, and delimited with special characters. The third is the "decoy" that only few users would pick. With this choice architecture, users are nudged towards the target password.

This technical report was **not** peer-reviewed and was published to complement a peer-reviewed publication at EuroUSEC '16.
If you wish to reference this paper, please cite it as: Tobias Seitz, *"The Decoy Effect for Passwords – A First Exploration"*, Ludwig-Maximilians-Universität München, Munich, Technical Report, July 2016, 8 pages.

In such a case, the participants generally favored option (A), the competitor. The difference in quality rating did not seem to make up for the higher price of option (B), the target. However, when a third option was introduced, the participants' preference changed:

Option	**Price**	**Quality rating** (0=worst, 100=best)
(A)	$1.50	50
(B)	$2.50	70
(C)	$3.00	60

Here, option (C) is the decoy. It serves as the least favorable option, because the beer is most expensive while rating even lower on the quality scale than option (B). The result is an asymmetry. Therefore, people are expected to choose option (B) more often because they can get a higher quality than with (C) for a lower price. The difference in quality between (A) and (B) now seems more graspable than before, while the crucial comparison is between (B) and (C). The decoy can be constructed in numerous ways by varying its values along the two dimensions, e.g. price and quality, as described in [2]. Ariely and Wallsten argue that people try to simplify the decision process using heuristics [3]. The fundamental workings of the effect lies in comparing the goods instead of evaluating them separately. Having the decoy item biases the comparison and generates a measurable effect on people's preferences. This kind of framing effect [4], is often called "choice architecture" [5]. Choice architectures have received recent attention by the usable security and privacy community (e.g. [6], [7], [8], [9]).



② Choose storage

How much storage is right for you?

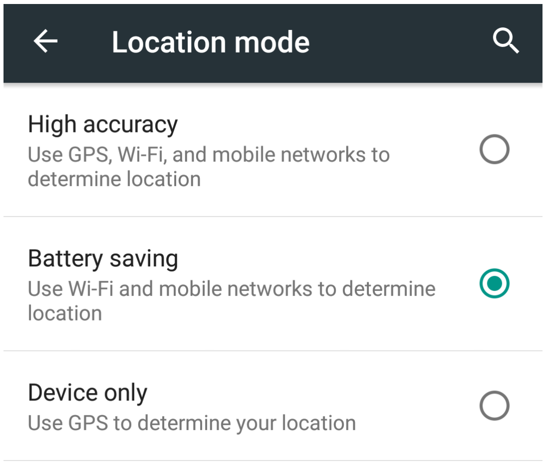
16GB[1] From $499	64GB[1] From $599	128GB[1] From $699

Fig. 2. The configurations in this online shop show a decoy pattern. Compared to the 16GB option, customers receive *four times* the storage by spending $100 more. Another $100 only *doubles* the storage from there. Hence, we see the 64GB version as the vendor's target, but other interpretations are possible. In any case, the mere juxtaposition instantly makes buyers compare and evaluate.

B. The Decoy Effect in Online Shops

Many companies use an elaborate choice architecture when they offer different configurations of the same product or different service levels. Figure 2 illustrates what this usually looks like in practice. One of the options commonly seems to be the "most reasonable". Often it is the one lying in the middle of two dimensions, e.g. price and storage capacity. The key is to get customers to compare the options while inconsistently increasing one dimension for the decoy product.

C. Decoys in Android



High accuracy
Use GPS, Wi-Fi, and mobile networks to determine location ○

Battery saving
Use Wi-Fi and mobile networks to determine location ●

Device only
Use GPS to determine your location ○

Fig. 3. Device location settings in Android 5. The "Device only" lacks information about accuracy and battery consumption and is thus least favorable. The hypothesized target seems to be the "Energy saving" mode, which the Android creators use to continuously collect anonymous location data.

We ask the question whether the device location settings in Android [1] have been constructed in a decoy architecture. We can identify the dimensions **accuracy** and **battery consumption** from the options (see Fig. 3). The "High accuracy" mode uses all available resources to quickly and accurately determine the device location using assisted GPS [10]. However enabling the GPS antenna drains the battery, so an additional "Battery saving" mode is available where GPS is disabled and position is only based on triangulation respectively multilateration, which makes it often less accurate. Still, the battery drain of having WiFi and the mobile network permanently enabled is noteworthy. The third and last mode is a "Device only" mode. It only uses GPS to determine the device location and thus this process involves a "cold start" where the device initially lacks information about satellites or device location in general. The cold start consumes a little more battery than with assisted GPS, but the accuracy outcome is comparable to the "High accuracy" mode. Nonetheless, if users do not disable their WiFi or data network, the battery drain of the "Device only" mode is probably similar to the "High accuracy" mode.

Thus we rank the accuracy and battery consumption[2]:

Mode	**Battery Drain**	**Accuracy**
High accuracy	High	High
Battery saving	Medium	Medium
Device only	Medium	High

If one tries to grasp the design choices for this architecture, at least one aspect is noteworthy: The "Device only" mode lacks information about both the accuracy and battery consumption, whereas those are clearly pointed out in the other two modes. From our ranking we can also see that in terms of battery consumption this may be a better trade-off than the "High accuracy" mode. For non-technical users, this may however be the decoy option. One could assume that the vendor would like to nudge users towards one of the upper two alternatives. A probable reason is the benefit for the creators of Android that those options provide: Users can only use these if they agree to provide continuous anonymous location data to the vendor. If they disagree, they are "stuck" with the "Device only" mode, and laymen might be very much attracted by the highlighted benefits of either high accuracy or low battery

[1] Android 6 is the latest public version at the time being.

[2] The ranks are only rough estimations to illustrate the effect.

consumption. Hence, the nudging strategy including the decoy pattern was likely not chosen by accident.

D. Passphrases and High Complexity

The decoy effect can only be applied when a person chooses between a set of alternatives. The alternatives need to be easily distinguishable regarding two dimensions. In the example above, the dimensions were price and quality. For passwords, we considered the dimensions **complexity** and **strength** suitable (see Fig. **??**). Thus, for the decoy effect these two dimensions need to be easily identifiable. For that matter, we opted to suggest passphrases instead of passwords. The chunks in passphrases are regular dictionary words, which makes the differences of multiple passphrases more evident [11].

From a usable security perspective, Shay et al. showed that word-based passphrases performed similarly to more shorter, yet more complex passwords [12]. When users create their passwords, a word-based composition policy has revealed benefits in terms of password strength [13]. In an online study, Shay et al. showed that a policy requiring passwords composed of two chunks with a total length of 16 characters (2word16) lead to stronger passwords than a more complex policy demanding fewer characters but more character classes (e.g. comp8 or 3class12). Keith et al. also showed that users' perception and ability to memorize passphrases depends on whether or not it contains regular punctuation [11]. Passphrases containing delimiters that we encounter in text processing were perceived as enjoyable.

In conclusion, we assumed that word-based passphrases were apt to demonstrate the decoy effect about the attractiveness of different passwords.

E. Non-Verbal Persuasion for Stronger Passwords

Our goal is to utilize password suggestions and the decoy effect to nudge users towards stronger passwords. Nudging users towards stronger passwords has been under constant research for years. For example, proactive password meters are well established and provide visual, non-verbal information about the entered password [14]. They are effective because they can persuade users try and achieve a high "score". Apart from the issue that the feedback provided is highly inconsistent across different services [15], it was also found that the way users try and increase their score is predictable [16]. For example, when they enter their usual password they might notice a rather low score from the password meter. The predictable reaction is to add numbers and/or an exclamation mark at the end, which then does not boost the strength significantly. We also address this issue in our concept by showing quality ratings as password meters (cf. Section III). Users can compare the strength of their self-chosen password to at least one alternative. This way they can be shown how to improve the rating in a more subtle manner. We hypothesize that instead of just adding a digit at the end of their re-used password, users might consider inserting an entire word or substitute a letter to reduce predictability.

Finally, we consider password suggestions *persuasive*. Weirich and Sasse were probably the first ones to put forward the understanding that users can be persuaded to alter their password behavior [17]. Other seminal work in persuasion for text-based passwords was done by Forget et al. [18], [19]. Like us, they mostly utilized suggestions to improve users' passwords. However, rather than suggesting a full password, their approach was denoted by modifying existing passwords. They found that suggestions are effective in increasing password strengths in regard to cracking attacks.

III. Choice Architecture for Decoy Passwords

To make use of the decoy effect to nudge users towards slightly stronger passwords, the complexity of the passwords should be easily comparable. Thus, we use passphrases that originate from regular dictionary words, as we think being able to read the passwords facilitates this comparison. All of the options we present to the users consist of four words plus alterations.

The **competitor** is a four word passphrase whose words are capitalized. There is no space or delimiter between the words. It receives a strength rating of (good). Trying to estimate the entropy of the competitor, we assume that the words are taken from a dictionary including 5000 words. Thus, there is a possible space of 5000^4 words, which translates to approximately 50 bits. Capitalization does not add to the entropy, since all words are capitalized uniformly.

The **target** contains random delimiters between the four words and a random special character at the end. The delimiters add approximately 3 bits of entropy, and the final special character adds another 3 bits. Thus, the total entropy of the target is 56 bits.

The **decoy** consists of four words add terminates in three random digits and a random special character. Furthermore, its characters are randomly capitalized. The random digits add 10 bits of entropy, the special character adds 3. The random capitalization is more difficult to estimate since it depends on the length of the generated password. Assuming a length of 25 characters, each character can either be lower- or uppercase, which results in 25 bits of entropy. Thus, we can estimate that the entropy of the decoy is around 88 bits.

We point out that even the competitor has an expected entropy that is likely higher than most real-world passwords. Yet, we were interested to see if the comparison of passphrases is suitable for the decoy effect.

IV. Research Goals

Since empiric evidence about the existence of the decoy effect in the realm of passwords is missing, our goal was to collect such evidence. We thus tried to answer the research question: Does preference for a password shift among users if a decoy is present, respectively does the effect exist at all?

V. Online Survey

The primary goal of our online survey was to answer our research question and find evidence for the existence of the decoy effect with passwords. If there was evidence in favor of the existence of the decoy effect, it would allow us to implement future concept around it.

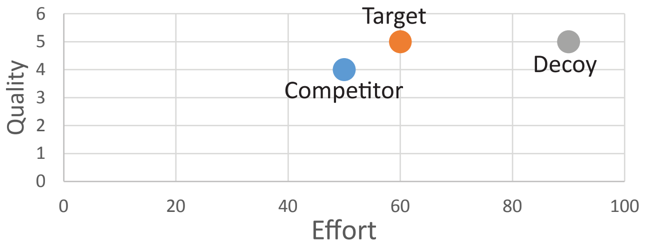


Fig. 4. The placement of options for the survey. The decoy does not bring any denoted benefit in terms of strength, while the effort to type and memorize it increases. In this configuration, the target and competitor are closer together than in our final concept.

A. Methodology

Since we were mostly interested in hypothetical preference and the decoy effect, a working prototype of the password generator seemed unnecessary at this point. Thus, we opted to show screenshots of certain sites overlaid by a password generator mockup. The survey consisted of a brief introduction to the task, voting for a passphrase for 12 different sites, a behavior and attitude assessment inspired by the questionnaires in [20] and the possibility to provide any kind of feedback on the study.

The generated passwords followed the choice architecture presented in Section III:

Generated password	**Quality**
(A) `BlueHouseNewStigma`	(good)
(B) `Blue.House.New.Stigma?`	(very good)
(C) `MirRorCloTHaTteNtioNCaBLe169#`	(very good)

Figure 4 depicts our intended quality-effort trade-off.

1) Study Design: The survey was executed in a between groups design. The independent variable was the presence of the decoy passphrase: for **group A** the decoy was missing, while it was shown for **group B**. We diversified the screenshots shown next to the password generator mock ups to make the participants evaluate their preference also based on the usage scenario. Moreover, we used different delimiters for the target passphrase to convey their random generation. Table I shows the details of these variations. To avoid ordering effects, the screenshots were shown in random order.

The twelve websites that served for the screenshots were selected bearing in mind that users make different efforts to create stronger passwords depending on the service they use it for [14], [21]. People categorize their accounts by their perceived value and vulnerability, but the assessment does not always base on hard facts. For example, Stobert and Biddle report that some participants in their qualitative study explained to use stronger passwords for sites they use frequently and vice versa [21]. Therefore, we tried to address the categories as outlined in Table I.

Moreover, to separate the chunks in the passphrase we only used special characters that do not require hitting the shift key on a standard, localized keyboard. In other words, these symbols are most frequently used in regular word processing which has been shown to be beneficial for passphrase design [11]. This benefit would only become apparent, though, if one were to type the passphrase oneself, which was not the case.

TABLE I. SCREENSHOTS OF THESE SITES SERVED AS BACKGROUND IMAGE FOR THE PASSWORD GENERATOR MOCKUPS IN THE ONLINE SURVEY. WE DIVERSIFIED ACROSS SIX CATEGORIES AND USED SIX DIFFERENT CHARACTERS THAT SERVED AS DELIMITERS.

Site Name	URL	Category	Delimiter
Postbank	postbank.de	Banking	-
Sparkasse	sskm.de	Banking	+
GMX	registrierung.gmx.net	Email	#
Google	accounts.google.com/SignUp	Email	.
Spiegel	spiegel.de/meinspiegel/register.html	News	,
Süddeutsche	id.sueddeutsche.de/register	News	.
Amazon	amazon.de/ap/register/[...]	Shopping	[space]
Zalando	zalando.de/login/	Shopping	-
Facebook	facebook.com	Social	+
Twitter	twitter.com	Social	#
Notizheft	notizheft.ch/registrieren.html	Unfamiliar	,
Ribbl	my.ribbl.com/registrierung	Unfamiliar	[space]

2) Tasks: On the landing page, participants were instructed that in the following, they should imagine to sign up on a number of websites. An image depicted an annotated output of the password generator on top of a blurred website screenshot. The suggested passphrases were also blurred to keep participants from noticing a different amount of suggestions later. Furthermore, the instruction stated explicitly that if they did not like any of the suggestions, they should vote for the one that they would most likely still agree to use. We decided to do this, since our main objective was measuring attractiveness and preference for one or the other.

At each decision point, we instructed the participants to pick from a given set of two (group A) or three (group B) password suggestions. The instruction always was *"Imagine your browser offers a password generator that suggests a secure password in specific situations. The password generator suggests multiple passwords to you on this website. Which of the suggestions would you choose?"*. Note that the respondents did not have to type the password themselves for two reasons. First, qualitative feedback from a pilot run of the study made it clear that typing is too cumbersome to complete the survey. The participants complained about the repetitiveness of the tasks. Only 13 of 40 respondents finished the pilot survey after one week. Second, in a realistic setting, a password manager might fill in the credentials for the user after they signed up, thus the most critical part to judge the attractiveness is the decision itself. Therefore, we offered radio buttons to select from the suggestions.

B. Hypotheses

We formulated the following hypotheses:

H0 If a decoy item is present, the preference for the target, respectively competitor passphrase will not be influenced. (Nullhypothesis)

H1 If the decoy passphrase is present, more people will vote for the target passphrase.

H2 If there is a preference for the *competitor* in the control group, then the *target* will receive more votes from participants in the experimental group.

1) Ethical considerations.: Our institution does not have an independent ethics committee for this kind of studies, but we endeavored to follow best ethical standards. Since responses in our online questionnaire were pseudonymous, immediate debriefing was difficult. However, after the study, we sent another email to all participants to let them know that a raffle had taken place among them. It also included a short debriefing, which explained the experimental design and a short disclaimer on secure passwords.

C. Recruiting and Demography

We leveraged an email distribution service at our university to invite 4719 students. We explicitly excluded students at the computer science, mathematics, statistics, and physics department as our concept primarily targeted mainstream users and we hoped to find those in the other departments. For example, Mazurek et al. found that students in business studies selected the most guessable passwords [22]. A raffle of three shopping vouchers of 10€ each incentivized participation in our survey. We announced a one-week deadline to participate.

180 people completed the questionnaire within a week, another 44 did not finish and were therefore not considered in the analysis. The random assignment to either group yielded $n = 88$ participants in group A, and $n = 92$ in group B. Thus, we achieved a near-optimally counterbalanced distribution in this regard. However, the gender distribution showed a strong skew towards female participants: 121 (67%) were female, 53 (29%) were male and 6 preferred not to answer. The participants' age was 25 years in average ($SD = 7.69$) and ranged from 17 to 68 years. We did not require respondents to state their study program, but many did nonetheless. The top five domains were humanities ($n = 32,\ 18\%$), medicine & pharmaceutics ($n = 29,\ 16\%$), natural sciences ($n = 19,\ 11\%$), teacher training programs ($n = 17,\ 9\%$), and economics ($n = 14,\ 8\%$).

D. Results

In total, we collected $N = 2160$ decisions, from which $n = 1056$ originated from group A, and $n = 1104$ from group B. We report test statistics on a significance level of $\alpha = 0.5$.

1) Score Calculation: In a first step, we calculated numerical scores for each participant, based on the number of their decisions for each option: The **T-score** for the number of times the target was chosen, and likewise the **C-score** for the competitor. These scores are available in both groups, while there was a third **D-score** in group B, where we counted the votes for the decoy password.

Additionally, the **T-C score** is the difference between decisions in favor of the target and the competitor password. This score can be calculated for responses in both groups. On the other hand, the **T-C-D score** only has an effect in group B. It is the number of decisions in favor of the target password by all other votes, including the decoy.

Finally, we established nominal categories depending on the T-C score. These categories represent the overall preference for one of the options. If a participant's T-C score was positive, we put them in the "**T_pref**" category. If it was negative we put them into the "**C_pref**" category. We labeled T-C scores of 0 with "**Indifferent**". These categories would allow us to run chi-squared tests to explore overall differences across groups.

One tricky aspect in examining preference is the different amount of options in the two groups. The advantage of our approach is that this factor is ruled out, without having to exclude responses: both the classification and the score calculation account for the different number of options.

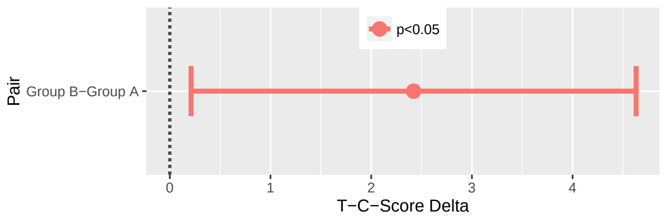


Fig. 5. The T-C score was significantly higher in group B of the online survey. Plot shows a confidence interval not overlapping the null hypothesis (zero difference).

2) Numerical Analysis: The average T-scores in both groups are very similar, $M_A = 7.03(SD_A = 4.02), M_B = 7.42(SD_B = 4.21)$. This results from a substantial part of group A favoring the target password in most scenarios, just like the participants in group B. A two-sample Wilcoxon rank-sum test revealed no significant differences regarding the T-score ($W = 3777.5, p > 0.05$).

Contrarily, there were significant differences regarding the C-scores ($M_A = 4.96(SD_A = 4.02), M_B = 2.93(SD_B = 3.48), W = 5269, p < 0.001, r = -0.26$) and also T-C scores ($M_A = 2.06(SD_A = 8.05), M_B = 4.48(SD_B = 6.94)$. Under the assumption that the T-C scores in group B are bigger, the effect is statistically significant, albeit relatively small ($W = 3381.5, p < 0.05, r = -0.16$). Figure 5 shows the confidence interval for the T-C score comparison.

The target password consistently received more votes from group B in each of the 12 voting tasks. Participants in group B chose the target password in significantly more test cases when they assessed the passwords with screenshots of Amazon ($\chi^2(1) = 12.16, p < 0.001$), Ribbl ($\chi^2(1) = 10.71,\ p < 0.01$), Sparkasse ($\chi^2(1) = 5.18,\ p < 0.05$), and Twitter ($\chi^2(1) = 4.18,\ p < 0.05$).

3) Category Analysis: The distribution of nominal preference is shown in Table II. A chi-squared test revealed a significant main effect across groups ($\chi^2(2) = 6.81,\ p < 0.05$ (two-tailed)). This indicates that the presence of the decoy boosted the target's attractiveness.

TABLE II. ABSOLUTE DISTRIBUTION OF RESPONDENTS ACROSS THE THREE CATEGORIES. SIGNIFICANTLY MORE PEOPLE IN GROUP B PREFERRED THE TARGET OVER THE COMPETITOR.

Group	C_Pref	T_Pref	Indifferent
A (Control)	31	50	7
B (Experimental)	17	63	12

4) Reversal of Preferences: In the case of Amazon and Ribbl, the preferences were reversed: While the group A revealed a clear preference for the competitor, group B preferred the target passphrase (see Figure 6). The target passphrase in these cases contained spaces as delimiters (cf. Table I). Since none of the other ten target passphrases contained spaces, we attribute the choice reversal to this special character.

5) Attitudes and Self-Reporting: The final part of the survey inquired on the impression about the suggested

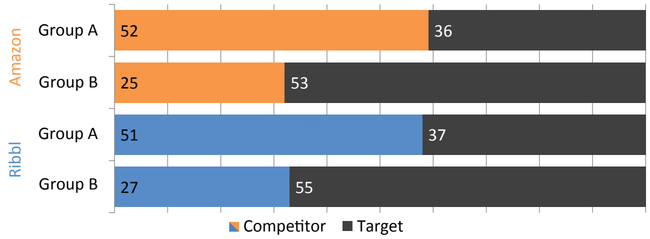


Fig. 6. Choice distribution for the two scenarios in which the target passphrase contained spaces as delimiters. In both cases, participants' preferences reversed depending on the presence of a decoy password (votes for the decoy not shown in diagram).

passphrases. Figure 7 shows the respondents' assessments. The willingness to use such a generator was generally low, despite the positive assessment of security and memorability of the suggestions.

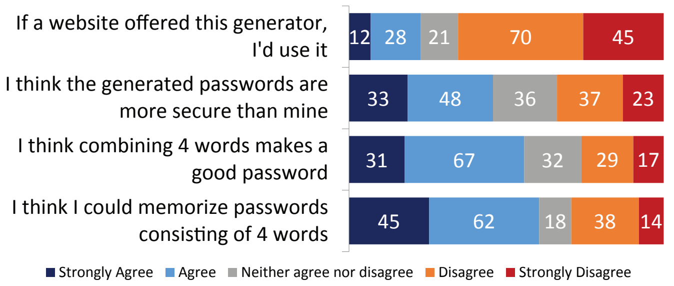


Fig. 7. Feedback on the concept from the participants in the online studies, measured on 5-point scales.

E. Result Summary

We could observe that the target password was favored by a large part of respondents, regardless of the study group. However, significantly more people voted for the target if the decoy was present. This effect is visible in the T-C score, but not the T-C-D score. We conclude that this supports **H1**. In the individual test cases, the decoy password only had a significant effect in 4 out of 12 scenarios. While in two of these scenarios the control group favored the competitor, the decoy group preferred the target password. The result is the hypothesized reversal of preference, but the evidence is still too little. Hence, we reject **H2** at this point.

Since our main hypothesis H1 was confirmed, we had reason to believe that the addition of the decoy made the target more salient if one merely has to choose. However, with the state of the choice architecture, we did not achieve the asymmetric dominance we initially intended, so we needed to improve it in a second step. Also the design of the survey and in particular the external validity of the results required further improvement, which we addressed in an online experiment. These and other results are shown in an upcoming paper at EuroUSEC '16.

VI. Discussion and Implications

We briefly discuss what we can learn from the results of our survey.

A. Decoys Influence Preference

To answer our initial research question: The addition of the decoy seems to have a measurable effect on the perceived attractiveness and or strength of passwords. At this point, the answer is preliminary and only remotely valid for the scenario in question. We can not yet confirm a real-world effect on self-selected passwords. However, we can assume that utilizing decoy passwords in systems where system-assigned passwords are mandatory is feasible. The decoy password might help to convince users of a password that provides a reasonable effort-strength trade-off. Thus, it is helpful to guide users to the more secure option while allowing a higher degree of autonomy [23]. In turn, users may be a little happier with their choice and take a more positive stand towards system-assigned passwords.

B. Strength Indicators

In our scenario the strength of the passwords was represented by a text-based label. We could see that most respondents preferred the passwords that received a better quality rating. Thus, one may assume that users only pick what they are told is better. However, the decoy password did not receive a better quality rating (although in theory it should because of its higher entropy) and did not attract as many votes as the target. Thus, we can assume that it was not the quality rating alone that lead to the preference shift, but also the composition of the password.

C. Decoys in Marketing vs. Passwords

We realized there may be fundamental differences between the way the decoy effect is used in consumer psychology and how we used it to influence password selection. For example, consumers actively seek an attractive product which they eventually want to buy. People usually do not go out to "select a password". The task is intrinsically unattractive for most of us. Thus, the usage scenario in passwords might not be optimal, as there is little evidence the decoy effect works for "unpleasant" decisions.

VII. Limitations

Our online survey was conducted as an initial exploration and it naturally shows limitations. First, our exploration was about preference only, so the assumptions might not hold true if we conduct the study in a more realistic environment. We also do not know with certainty if the participants believed the quality ratings of the passwords: We could see that they preferred the "better" passwords, but we did not receive qualitative feedback on why they made this assessment. Moreover, the quality rating of the passwords was very conservative. In reality, we expect that all three suggested passwords would withstand a high number of guesses.

We also did not fully counterbalance the websites with the delimiters. However, the main focus lied on the aggregated preference and we expect that a fully counterbalanced design would not have narrowed confidence intervals. Finally, for the same reason, we did not measure memorability effects at this point, which should be done in future work.

VIII. Conclusion and Future Work

We presented a first exploration of the decoy effect and a potential application in usable security and privacy. Our use case is a concept for persuasive password generation relying on the central idea to suggest three generated phrases and compose them with regular words. The user study showed that a large part of the selections in our choice architecture is predictable. Yet in that scenario, this is a good thing: Participants opted for a stronger password if the stronger options outnumbered weaker suggestions. Although we have found cues that preferences for specific passphrases are reversible, we are cautious about attributing all observations to the decoy effect.

Future research should also investigate additional qualities of password suggestions. We see great potential to learn the user's composition strategy and adapt suggestions to make them more attractive and effective. The concepts should be made available to the public in a real-world deployment. This will also allow to collect data in the field and address the limitations of our study.

In conclusion, there may be use cases where the decoy effect will prove beneficial in usable security. Users face many choices when the decide for privacy settings (cf. Section II-C) or application permissions, or an authentication scheme in general [24]. Nudging with decoys may work quite well in such scenarios, which poses immediate research questions that are yet to be answered.

Disclaimer

This paper was **not** peer reviewed. Although great care was taken to compile this work, some statements express the author's personal opinion and should not be referenced without prior validation.

Acknowledgments

The author would like to thank Isabel Schönewald for her preliminary studies, and also Emanuel von Zezschwitz, Heinrich Hussmann and the anonymous reviewers for CHI'16 and EuroUSEC '16 for their input on previous versions of the paper.

References

[1] B. Ur, J. Bees, S. M. Segreti, L. Bauer, N. Christin, and L. F. Cranor, "Do users' perceptions of password security match reality?" in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, ser. CHI '16. New York, NY, USA: ACM, 2016, pp. 3748–3760. [Online]. Available: http://doi.acm.org/10.1145/2858036.2858546

[2] J. Huber, J. W. Payne, and C. Puto, "Adding Asymmetrically Dominated Alternatives: Violations of Regularity and the Similarity Hypothesis," *Journal of Consumer Research*, vol. 9, no. 1, p. 90, 1982.

[3] D. Ariely and T. S. Wallsten, "Seeking Subjective Dominance in Multidimensional Space: An Explanation of the Asymmetric Dominance Effect," *Organizational Behavior and Human Decision Processes*, vol. 63, no. 3, pp. 223–232, 1995. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0749597885710758

[4] D. Kahneman and A. Tversky, "Choices, Values, and Frames," *American Psychologist*, vol. 39, no. 4, pp. 341–350, 1984.

[5] R. H. Thaler, C. R. Sunstein, and J. P. Balz, "Choice architecture," *Social Science Research Network*, no. August, 2010. [Online]. Available: http://ssrn.com/abstract=1583509

[6] L. Coventry, P. Briggs, D. Jeske, and A. V. Moorsel, "SCENE : A Structured Means for Creating and Evaluating Behavioral Nudges in a Cyber Security Environment," in *Design, User Experience, and Usability. Theories, Methods, and Tools for Designing the User Experience*, 8517th ed., A. Marcus, Ed. Springer International Publishing, 2014, pp. 229–239.

[7] S. Egelman, A. P. Felt, and D. Wagner, "Choice Architecture and Smartphone Privacy: Theres A Price for That," in *The economics of information security and privacy*, R. Böhme, Ed. Springer, 2013, pp. 211–236.

[8] A. Jameson, S. Gabrielli, P. O. Kristensson, K. Reinecke, F. Cena, C. Gena, and F. Vernero, "How can we support users' preferential choice?" *Proceedings of the 2011 annual conference extended abstracts on Human factors in computing systems - CHI EA '11*, p. 409, 2011. [Online]. Available: http://portal.acm.org/citation.cfm?doid=1979742.1979620

[9] S. Korff and R. Böhme, "Too Much Choice: End-User Privacy Decisions in the Context of Choice Proliferation," in *Symposium on Usable Privacy and Security (SOUPS '14)*, 2014, pp. 69–87. [Online]. Available: https://www.usenix.org/system/files/soups14-paper-korff.pdf

[10] G. M. Djuknic and R. E. Richton, "Geolocation and Assisted GPS," *Computer*, vol. 34, no. 2, pp. 123–125, 2001.

[11] M. Keith, B. Shao, and P. Steinbart, "A Behavioral Analysis of Passphrase Design and Effectiveness," *Journal of the Association for Information Systems*, vol. 10, no. 2, pp. 63–89, 2009. [Online]. Available: http://aisel.aisnet.org/cgi/viewcontent.cgi?article=1492&context=jais

[12] R. Shay, P. G. Kelley, S. Komanduri, M. L. Mazurek, B. Ur, T. Vidas, L. Bauer, N. Christin, and L. F. Cranor, "Correct Horse Battery Staple," in *Proceedings of the Eighth Symposium on Usable Privacy and Security (SOUPS '12)*. New York, NY, USA: ACM, 2012, pp. 1–20. [Online]. Available: http://dl.acm.org/citation.cfm?doid=2335356.2335366

[13] R. Shay, S. Komanduri, A. L. Durity, P. S. Huh, M. L. Mazurek, S. M. Segreti, B. Ur, L. Bauer, N. Christin, and L. F. Cranor, "Can Long Passwords Be Secure and Usable?" in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*, 2014.

[14] S. Egelman, A. Sotirakopoulos, I. Muslukhov, K. Beznosov, and C. 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 (CHI '13)*, 2013, pp. 2379–2388. [Online]. Available: http://doi.acm.org/10.1145/2470654.2481329

[15] B. Ur, P. G. Kelley, S. Komanduri, J. Lee, M. Maass, M. L. Mazurek, T. Passaro, R. Shay, T. Vidas, L. Bauer, N. Christin, and L. F. Cranor, "How Does Your Password Measure Up? The Effect of Strength Meters on Password Creation Blase," in *Security'12 Proceedings of the 21st USENIX conference on Security symposium*, 2012, pp. 5–16. [Online]. Available: https://www.usenix.org/system/files/conference/usenixsecurity12/sec12-final209.pdf

[16] S. Komanduri, R. Shay, P. G. Kelley, M. L. Mazurek, L. Bauer, N. Christin, L. F. Cranor, and S. Egelman, "Of Passwords and People," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)*, 2011, pp. 2595–2604. [Online]. Available: http://dl.acm.org/citation.cfm?doid=1978942.1979321

[17] D. Weirich and M. A. Sasse, "Pretty Good Persuasion: A First Step towards Effective Password Security in the Real World," in *Proceedings of the 2001 Workshop on New Security Paradigms (NSPW '01)*. New York, NY, USA: ACM, 2001, pp. 137–143. [Online]. Available: http://dl.acm.org/citation.cfm?id=508195

[18] A. Forget, S. Chiasson, P. C. Van Oorschot, and R. Biddle, "Improving Text Passwords Through Persuasion," in *Proceedings of the 4th Symposium on Usable Privacy and Security (SOUPS '08)*. New York, NY, USA: ACM, 2008, pp. 1–12. [Online]. Available: http://portal.acm.org/citation.cfm?id=1408666

[19] A. Forget and R. Biddle, "Memorability of Persuasive Passwords," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '08)*, 2008, p. 3759. [Online]. Available: http://portal.acm.org/citation.cfm?doid=1358628.1358926

[20] C. Kuo, S. Romanosky, and L. F. Cranor, "Human Selection of Mnemonic Phrase-Based Passwords," in *Proceedings of the second*

Symposium on Usable Privacy and Security (SOUPS '06), 2006, pp. 67–78. [Online]. Available: http://doi.acm.org/10.1145/1143120.1143129

[21] E. Stobert and R. Biddle, "The Password Life Cycle: User Behaviour in Managing Passwords," in *Proceedings of the 10th Symposium On Usable Privacy and Security (SOUPS '14)*. New York, NY, USA: ACM, 2014, pp. 243–255.

[22] M. L. Mazurek, S. Komanduri, T. Vidas, L. Bauer, N. Christin, L. F. Cranor, P. G. Kelley, R. Shay, and B. Ur, "Measuring password guessability for an entire university," in *Proceedings of the 2013 ACM SIGSAC conference on Computer & communications security - CCS '13*, 2013, pp. 173–186. [Online]. Available: http://dl.acm.org/citation.cfm?doid=2508859.2516726

[23] R. M. Ryan and E. L. Deci, "Self-Determination Theory and the Facilitation of Intrinsic Motivation," *American Psychologist*, vol. 55, no. 1, pp. 68–78, 2000.

[24] A. Forget, S. Chiasson, and R. Biddle, "Choose Your Own Authentication," in *Proceedings of the New Security Paradigms Workshop (NSPW '15)*. Twente, The Netherlands: ACM, 2015.