Why Do They Do What They Do?: A Study of What Motivates Users to (Not) Follow Computer Security Advice

Michael Fagan and Mohammad Maifi Hasan Khan, ***University of Connecticut***

https://www.usenix.org/conference/soups2016/technical-sessions/presentation/fagan

This paper is included in the Proceedings of the Twelfth Symposium on Usable Privacy and Security (SOUPS 2016).

June 22–24, 2016 • Denver, CO, USA

ISBN 978-1-931971-31-7

Open access to the Proceedings of the Twelfth Symposium on Usable Privacy and Security (SOUPS 2016) is sponsored by USENIX.

Why Do They Do What They Do?

A Study of What Motivates Users to (Not) Follow Computer Security Advice

Michael Fagan
University of Connecticut
Storrs, Connecticut, USA 06269
michael.fagan@uconn.edu

Mohammad Maifi Hasan Khan
University of Connecticut
Storrs, Connecticut, USA 06269
maifi.khan@uconn.edu

ABSTRACT

Usable security researchers have long been interested in what users do to keep their devices and data safe and how that compares to recommendations. Additionally, experts have long debated and studied the psychological underpinnings and motivations for users to do what they do, especially when such behavior is seen as risky, at least to experts. This study investigates user motivations through a survey conducted on Mechanical Turk, which resulted in responses from 290 participants. We use a rational decision model to guide our design, as well as current thought on human motivation in general and in the realm of computer security. Through quantitative and qualitative analysis, we identify key gaps in perception between those who follow common security advice (i.e., update software, use a password manager, use 2FA, change passwords) and those who do not and help explain participants' motivations behind their decisions. Additionally, we find that social considerations are trumped by individualized rationales.

1. INTRODUCTION

Academics have widely accepted that privacy is not only valued by individuals, but also helps aspects of our society function [24]. Computer/data privacy is no different: many report putting a high value on the ability to control who can access their data and information [17, 22]. Since security of computers is the first step towards computer privacy, it is imperative that we not only create new, stronger cryptographic and security tools, but that we also understand how to best motivate users to adopt new tools and techniques.

The facts of what people can do to stay safe and how they use that advice have been well studied, with many finding divergence between recommended and actual protections [15, 8]. The failure of current and past motivational and/or security approaches [5, 1], lack of information about many facets of the problem, including adaptability of many security advices [12, 13], and issue specific (e.g., updating) concerns [29, 28, 11] have all been noted as part of the explanation for the gap. That said, to the best of the authors' knowledge, no one has broadly approached the question of **"why do some follow security advice, while others do not,"** using empirical data collected and analyzed for that purpose. Though some work has looked at the concerns of users in many specific scenarios of user security, we seek to sift through the context-specifics and overall economics of some security decisions, with the hopes of gaining insight into the overall problem. By investigating these kinds of trends, we can better understand motivation in this area, and evaluate/improve current approaches towards increasing the security of users.

With this study, we investigate the motivations of users to follow or not follow common computer security advice. We model decision-making with a rational, cost/benefit framework, expanding it to include both the concept of risk, which is expected to be key to security decisions, as well as social motivations. Using this grounding in interdisciplinary prior work, we design a web-based survey distributed to those 18 and over living in the U.S. via the service Mechanical Turk. We use 4 common security recommendations (i.e., updating software, use of a password manager, use of 2FA, changing passwords) as a foundation for our surveys. With each advice, we form two groups: one of those who follow the advice (Yes groups), and one that does not (No groups). In all, we collect 290 survey responses constituting both qualitative and quantitative data. Through analysis of this data, we extract the following key findings related to the question "why do some follow security advice, while others do not?":

- Benefits of following are rated higher by those who follow each advice compared to those who do not. Those who do not follow rate the benefits of doing so as higher than the groups that practiced each advice.
- Risks of not following are rated higher by those who follow each advice compared to those who do not.
- Costs of not following are also seen as higher by those who follow each advice compared to those who do not for all but one case (using 2FA).
- Security and convenience are common themes in the qualitative comments. For all tools, Yes groups in many cases report following because they think doing so is more secure. In some cases (i.e., updating and using a password manager), those who follow are also drawn by added convenience.
- Individual concerns are rated higher than social concerns for all variables, indicating low social motivations around computer security.

Symposium on Usable Privacy and Security (SOUPS) 2016, June 22–24, 2016, Denver, Colorado.

These highlighted findings and the full results presented in this paper help towards understanding why those who follow computer security advice, do and those who do not follow, don't. This study continues a long running track approaching this problem and brings new information to the debate on how to best address it. We find evidence to support prior suggestions that users know the costs/benefits involved, but also see gaps between those who do and do not follow each advice on benefits and risks. This implies that at least one or both of the groups for each advice are miscalculating in their considerations, since both positions cannot be simultaneously "true." Knowing who is truly "wrong" is imperative, and so the authors echo the calls of prior work [12] on the need for more data about what users think and experience, as well as measures of actual risk to best interpret ours and other's results.

2. RELATED WORK

This study is influenced by several arms of usable security research. First, our work supposes that for all security related decisions, users are making a rational choice by weighing the costs against the benefits. Second, we add perception of risk to our considerations since this is integral to motivation around secure behavior. Third, we argue there is or should be a social component to users' motivations, so that aspect is also incorporated into this study. Fourth, we choose to look at the dichotomy between those who adhere to good security behavior and those who do not. All four of these tenets are grounded in the literature.

2.1 Security Decisions as a Rational Choice

Though complex, human decision-making can be viewed as a consideration of costs and benefits, where humans are rational actors who choose to minimize cost and/or maximize benefit. This view of computer security decision-making has been prominent. Herely in 2009 was one of the first to suggest that users' failure to adhere to good security behavior could be attributed to them finding the costs too high and/or benefits too low [12]. He supports this supposition by citing the low chance of an actual security breach for any given user and the high cost of daily security maintenance. Herley goes on to suggest that more data is needed to determine the actual costs and benefits of these decisions to better inform the advice experts give. By 2014, Herley found that the approach of researchers had not changed much, leading him to say in a follow-up work:

> It is easy to fall into the trap of thinking that if we can find the right words of slogan we can convince people to spend more time on security. . . . We argue that this view is profoundly in error. It presupposes that users are wrong about the cost-benefit tradeoff of security measures, when the bulk of the evidence suggests the opposite.

What Herley suggests is that rather than users being ill-informed about security, they could be making a perfectly rational decision, at least in their eyes. This view is echoed by Geordie Stewart's 2012 work "Death by a Thousand Facts." Here, Stewart and David Lacey argue that "security-awareness" based approaches to increasing user security have and will continue to fail because, unlike how some researchers assume, users are not ignorant of good security behavior [5].

On the other hand, many studies have generated results that suggest users are miscalculating the costs and benefits. "Out of the Loop" by Rick Wash et al. found that a significant portion of sampled Windows 7 users did not understand what updates were changing in their system and could not execute their intentions for computer management [29]. Vaniea et al.'s "Betrayed by Updates" has similar findings that suggest prior negative past experience could play a large role in users deciding not to apply updates [28].

This divide in the literature on user motivations around computer security could be related to differences in perceptions between people about computer security. Specifically, it is possible that experts and others who follow advice do see the costs and/or benefits of adhering to good security behavior differently. Our study hopes to investigate this view of the issue to shed light on the motivations of everyone around these decisions, but we also extend the simple cost-benefit decision model for the context.

2.2 The Significance of Risk Perception

For security decisions, the literature shows us that risk perception, specifically a user's idea about the possible negative outcomes resulting from their decsions is key towards understanding security related behavior. Howe's 2012 survey of work about human psychology in computer security identified that security risks and risk perceptions were central considerations for many researchers [14]. Studies that have investigated mental models, such as Camp's 2006, Asgharpour et al.'s 2007, and Kang et al.'s 2015 works as well as other studies that looked directly at risk perceptions in different contexts, all focus on the importance of risk in the very design of their studies [4, 3, 16, 9, 11].

Some researchers have gone further and have tried to alter risk perceptions to improve communication and/or motivations. Harbach et al.'s 2014 work that appeared in CHI leveraged personal information to highlight the effect of Android permissions on user's data [10]. This was meant to alter their perception of the risks associated with each permission they are asked to grant, hopefully making them realize what exactly is at stake. The study found that users made more privacy-conscious decisions when presented with such information during app installation.

Since the perception of risk in particular has been repeatedly highlighted in work investigating security motivations, our study separates "cost" into explicit cost/inconvenience (e.g., time, money) and risk to provide a fuller picture of participants' perceptions and motivations.

2.3 Social Motivation

Though risk perception is intrinsically linked with security decisions, we also add another component absent from many other studies on this subject. Social motivations are integral towards voluntary compliance. Tyler's 2010 book "Why We Cooperate" details his theory on human motivation and cooperation [27]. In short, he argues that social motivations (i.e., motivations driven by values or wanting to help/please others) are much stronger and longer lasting than instrumental motivations (i.e., motivations related towards gaining material reward or avoiding material cost). Tyler presents his theory in contrast to the view of social motivations as simply a kind of instrumental motivation. Rather than trying to gain a future material benefit from

someone, Tyler says people who act in a socially positive way do so because they're motivated by their existing social connections.

The importance of social motivation is not new. Prosocial behaviors, as they are sometimes called have been studied for decades, being investigated all levels (i.e., individual, small, and large groups) and in many contexts [20].

Though Tyler and many others have theorized on the source of prosocial behavior and by extension social motivations, that debate is beyond the scope of this work. We simply accept that social motivations, regardless of the biological or psychological source, are important to human decision-making. Thus, our study considers participants' motivations with regards to other users, if any.

Ours is not the first work to acknowledge the importance of social considerations in technology decision-making. In SOUPS 2014, Das et al. found that social motivations play a role in cyber-security behavior [6]. Specifically, they found that observability of a secure behavior was a "key enabler of socially triggered behavior change," showing that social motivations could be important to technology decisions. The authors showed that users could be better motivated to act securely online if their peers would know the decisions they were making.

2.4 "Good" Actors and "Bad" Actors

Though Herley may be right and users may be properly assessing the computer security situation when they make what seems to be poor decisions, there is evidence in the literature suggesting that experts and average users do think and act differently when it comes to computer security. Two recent reports that support this statement appeared in SOUPS 2015. One, Ion et al.'s "No one can hack my mind..." showed that experts and regular users reported different behaviors when asked which they think are the best for staying safe, showing a divide in thinking [15]. The study also found that experts reported different security behaviors than non-experts. Additionally, another SOUPS 2015 work, "My Data Just Goes Everywhere" by Kang et al. found that mental models of computer security and privacy were different, specifically that average users had simpler models than expert users, again showing a difference in thinking between experts and everyone else [16]. The authors further found that more detailed models enabled experts to articulate more privacy threats, the first step towards avoiding them. That said, Kang also found that there was no direct correlation between participants' technical background and the actions they took to control their privacy, indicating that even those who should know better sometimes behave insecurely.

Though there are many documented differences between experts and average users, there is also substantial evidence that an expert is not necessarily a "good" actor. Our study wants to examine the difference in motivation between "good" and "bad" actors, which in this context are those who adhere to secure behavior and those who do not, respectively. As such, rather than compare experts with non-experts, our study compares those who report following common security advice with those who report not following such advice.

Combing all these concepts, the authors developed a web-study to gain insight into many aspects of why some users may follow computer security recommendations while others ignore them. The results of this study transcend prior work on this topic by collecting and analyzing a large dataset containing a sample of users' self-reported motivations for following or not following a broad range of security advice.

3. METHODS

Our study design incorporates both quantitative and qualitative methods to help outline differences between users on the topic of four instances of security advice, which are as follows:

1. Keeping your software up to date
2. Using a password manager
3. Using two-factor authentication
4. Changing passwords frequently

1-3 are commonly recommended by computer security experts to help users stay safe. Advice 4 is a common folk advice that isn't necessarily recommended by experts. All are extracted from Ion et al.'s 2015 work [15]. For each, we formulate two groups of users, one that follows (who uses the tool or does the action) and one that does not follow (does not use the tool or do the action). We are interested in comparing these groups because we want to understand why there is a decision gap between otherwise similar users, which would help towards identifying ways to encourage better online behavior among more of the Internet-using population. To help in describing the study, we will refer to samples of users who follow each advice as "Yes" groups for those respective advices, while we will refer to samples of users that do not follow each as "No" groups.

As explained in the prior section, we use a rational choice perspective to frame our study. Using cues from multiple recent works [4, 3, 16], we extend the traditional cost/benefit analysis to include perception of risk. Finally, we consider the social aspect of each decision as well, also inspired by recent literature [27, 6]. This study has 12 variables we investigate, named as follows:

1. *Individual Benefit of Following*
2. *Social Benefit of Following*
3. *Individual Cost/Inconvenience of Following*
4. *Social Cost/Inconvenience of Following*
5. *Individual Risk of Following*
6. *Social Risk of Following*
7. *Individual Benefit of Not Following*
8. *Social Benefit of Not Following*
9. *Individual Cost/Inconvenience of Not Following*
10. *Social Cost/Inconvenience of Not Following*
11. *Individual Risk of Not Following*
12. *Social Risk of Not Following*

Since Yes groups assumedly follow the advice and No groups report not following the advice, the same survey question phrasing could not be used to define each variable for both groups. Thus, we must compare responses to a slightly different question from each in our analysis. In other words, we must contrast what those who follow say their experience is to what those who do not follow expect their experience to be if they *did* follow. Specifically, variables 1-6 are defined using the following phrasings:

Yes How much would you say [you | users of other computers] are [benefited | cost or inconvenienced | put at risk] by you [following the advice]?

No How much would you say [you | users of other computers] would be [benefited | cost or inconvenienced | put at risk] if you did [follow the advice]?

Note that variables 1-6 are defined for Yes groups using the phrase highlighted above while the same variables are defined with the other phrase for No groups. The portion of the phrasings above that are separated by vertical bars and/or in brackets are the wordings used to form the question for each of the variables 1-6 we test, as appropriate. For example, "you" is used to replace the first bracket for *Individual* variables, while "users of other computers" is used for *Social* variables. The second brackets are likewise replaced for variables that ask about benefits, costs/inconveniences, and risks, respectively. Finally, "follow(ing) the advice" is replaced as appropriate for each advice that we test in the surveys (e.g., the 2FA Follow groups' was replaced with "using/use two-factor authentication," etc.).

Similarly, variables 7-12 are defined using the following phrasings:

Yes How much would you say [you | users of other computers] would be [benefited | cost or inconvenienced | put at risk] if you did not [follow the advice]?

No How much would you say [you | users of other computers] are [benefited | cost or inconvenienced | put at risk] by you not [following the advice]?

Instruments for these variables are created in the same fashion as described above.

As mentioned, each variable is defined in a slightly different format for Yes and No groups, meaning our analysis will compare ratings that are more/less hypothetical depending on the group. For example, considering the *Individual Benefit of Following*, Yes groups' reported benefits will be compared with the benefits No groups report they *would* get from following the advice. Though there may seem to be an issue with comparing hypothetical ratings to more grounded reports, the goal of this work is to identify the possible gaps in perceptions between those who follow security advice and those who do not. For many decisions, but particularly in the contexts examined in this study, a user must imagine at least some hypotheticals when considering whether to follow an advice or not, since the user may be pondering a behavior they have not practiced in the past. We hope to identify the skewed or biased perceptions users may have about the possible outcomes, thus it is valuable to compare the reported effects from followers of an advice with the projected effects from those who do not currently follow. This requires comparing some more hypothetical ratings with those that are more grounded.

Surveys containing the instruments described were created for each group (Yes/No) for each of the 4 tested pieces of advice. A qualitative question asking survey-takers *why* they chose to follow or not follow the target advice (i.e., "Please explain in a few sentences why you choose to (not) [follow the advice].") was also included. The qualitative question was shown first, alone on a separate page in all surveys to avoid biasing the open-ended responses towards our overall study framework as seen in the structure of other survey instruments (i.e., the focus on benefits/costs/risks). On the next survey page participants were asked about benefits/costs/risks of the actual target decision they reportedly made, followed by the benefits/costs/risks they *would* expect if they made the opposite of their decision on the final page. Survey templates for both groups, showing the order of questions, can be seen in the Appendix.

3.1 Sampling Methodology

Participants were recruited with a single Mechanical Turk posting that showed the information sheet for the study and directed interested users to a University-hosted Qualtrics survey that asked basic demographic questions and 4 screening questions to be used to assemble the Yes and No groups. Participants who responded to the screening survey were compensated $0.25 for their time and effort. The full screening survey can be seen in the Appendix.

After collecting the screening data, samples of 50 participants were assembled into 8 groups (one for the Yes and one for the No groups for each advice). Unique and independent Yes and No groups were formed by randomly selecting participants who reported "Yes" or "No," respectively to the screening question of whether or not they follow each advice.

Each group of participants was contacted with their corresponding group survey. Participants were contacted with a link through Mechanical Turk's messaging system that directed them to a new posting with the same information sheet as before, but this time a link to the appropriate survey. Participants were informed that they could take this survey if they wanted, but were under no obligation to reply. If they chose to answer, they were compensated another $4 for their time and effort on the longer survey.

3.2 Coding Methodology for Qualitative Data

To facilitate useful analysis of the qualitative data collected, we adopted a Grounded Theory approach to developing our codebook and coding our data [25]. The codebook was developed by the lead researcher, with the addition of some codes generated during analysis. Deductive codes, based on the study design and pertinent literature, along with the structure of the codebook were developed before data collection began. The focus here was on broad concepts like "avoid risk" or "increase security" since context specific codes could be best developed inductively, while looking at the data. There were seven deductive codes developed.

When data was collected, a random sample of one third of all comments from each group was selected and used to develop inductive codes by the lead researcher that focused on more specific concerns extracted from user comments. Some examples of codes developed though inductive coding are "I don't want to" and "increase financial security," showing the range of reasons given by participants. Since reasons between groups and advices varied, most of these codes were not broadly applicable, but some were. For example, "Low/no risk/Don't care if hacked" is an inductive code that was applied many times for several instances of advice. A total of 32 inductive codes were created for all groups. These codes (deductive + inductive) were used as

the codebook by another researcher who was less involved in the study and its design than the lead investigator. The same codebook was used to analyze all qualitative data. In the process of coding, several more tags were created from patterns found in the full samples, which were added to the codebook. Twenty-four of these codes were created.

Using the methodology described in this section, we collected a sample of active Internet users' motivations to follow or not follow computer security advice, which we analyze in the next section.

4. EVALUATION

To drive our analysis, we formulate the following hypotheses related the question "why do some follow security advice, while others do not?"

- H-1a For all decisions, the *Benefits of Following* will be seen as higher by the Yes groups compared to the No groups.
- H-1b For all decisions, the *Benefits of Not Following* will be seen as higher by the No groups compared to the Yes groups.
- H-2a For all decisions, the *Risks of Not Following* will be seen as higher by the Yes groups compared to the No groups.
- H-2b For some decisions, the *Risks of Following* will be seen as higher by the No groups compared to the Yes groups.
- H-3a For all decisions, the *Costs of Not Following* will be seen as higher by the Yes groups compared to the No groups.
- H-3b For all decisions, the *Costs of Following* will be seen as higher by the No groups compared to the Yes groups.
- H-4a Those who follow each advice will do so, generally, to increase their security and/or for convenience purposes.
- H-4b Those who do not follow each advice will do so, generally, to avoid a cost/inconvenience or due to confidence in current behavior (i.e., they might know they should change, but don't want to).
- H-5 *Social* considerations will be lower than *Individual* concerns for all decisions.

These predictions are based on prior findings and intuition. With hypotheses 1-3, we contend that participants will rate the benefits/costs/risks of their decision in a way that justifies their decision. For example, it is likely that each group will see the benefits gained and risks avoided by their decision as higher than if they had made the opposite. We expect the reasons given for these decisions will center on security and convenience, as these are the two core things at stake in many of these cases. Finally, the magnitude of social concerns is expected to be lower than individual concerns due to the nature of computing, which physically separates individuals, possibly obscuring how one's decisions affect others online. Please note, Hypothesis 2b is expected to only apply to some of the tested advice. This divergence compared to the other hypotheses is based on clues from data collected by the authors for prior studies about user decisions in the contexts of software updates, using 2FA, and using a password manager.

Before we demonstrate how these hypotheses are supported by our data, we first describe the overall sample collected and each group in terms of demographics. Once the make-up of our participants is established, we use our hypotheses to guide the rest of our evaluation.

4.1 Sample Details

As explained in Section 3, we collected an initial sample from Mechanical Turk using a short screening survey. A total of 805 participants enrolled in this step, but not all were considered for inclusion in groups to be contacted with follow-up surveys. We removed participants with incomplete answers on the screening survey and those who did not own a computer of their own from the eligible list, which reduced the pool to 764.

59% of the 764 are male, while 41% report female as their gender. The overall average age of participants is 34 years old. When asked how often they use the computer, 96% report "Often" or "All the time." The average general computer expertise rating is 4.15, while the average rating for computer security is 3.6. In both cases participants were simply asked "How would you rate your [general computer | computer security] expertise?" Both are measured on a 5-point Likert scale, anchored at 1-Very Poor and 5-Very Good. Though the instruments for measuring expertise are broadly defined, which could result in some level of error from a "true" measure, our approach was deliberate in order to ascertain the participants' general confidence in their computer and security knowledge and proficiency. The statistics are merely used to describe the sample collected and did not influence group forming or analysis other than attempting to control the variables between groups, where possible.

These statistics are not representative of the general population due to the nature of Mechanical Turk and the voluntary recruitment method used. That said, responses to the screening questions used for grouping show similar statistics as reported in prior studies [15]. Adoption rates for some advice were seemingly higher than would be expected for the general population, an effect that could also be attributed to the nature of the Mechanical Turk population or self-selection in the recruitment methods. Summaries of responses to grouping questions from the full sample of 764 can be seen in Table 1. In all, our sample represents a group of active computer users who generally rate their computer and security proficiency as higher than average, but are not all followers of the tested advice.

We formed 8 randomly selected, independent, unique groups of 50 participants each from the full sample initially collected. A participant is considered eligible for a group if they are not already in another group and exhibit the group's target behavior. For example, only participant who answered "Yes" to the question "Do you keep your computer's software up to date?" were considered eligible for the Yes group for the updating advice. The groups of 50 were gender-balanced so that 25 eligible males and 25 eligible females were contacted for each group. One group, those who do not keep their software up to date, only had 47 eligible participants out of the total pool of 764.

Not all participants contacted for each group responded. All groups ended up with 30-40 participants, which are used for this analysis. Details about the profile of each group sample can be seen in Table 2.

	Yes	No	I Don't Know
Do you keep your computer's software up to date?	701 (92%)	47 (6%)	15 (2%)
Do you use a password manager (e.g., LastPass, OnePass, KeePass) to manage your online account passwords?	157 (21%)	599 (78%)	8 (1%)
Do you use two-factor authentication (e.g., 2-Step Verification) for at least one of your online accounts?	471 (62%)	210 (28%)	81 (10%)
Do you change your passwords frequently?	311 (41%)	446 (58%)	5 (1%)

Table 1: Response frequencies (and rates) from all initial participants who own their own computer and completed the full screening survey (n=764) for each question used to form groups.

All group samples are similar on all demographic questions except self-rated computer security expertise, which had significant differences between groups when tested using a Kruskal-Wallis test [19]. Self-rated security expertise (i.e., "How would you rate your computer security expertise?") is lower for some No groups (i.e., update, changing passwords). Tests of the correlation between participants' rating for security expertise and their responses to survey instruments for all our variables using Spearman's correlation coefficient [2] resulted in no strongly significant values ($\forall$, $p > 0.05$, except *Individual Benefit of Following* where $p = 0.045$). This suggests that though there are slight differences between some Yes and No groups for self-rated security expertise, security expertise itself is not a good predictor of most perceptions. Essentially, as best as we can measure, the groups we compare are similar in most respects, security expertise being a notable exception, but even this difference is only apparent between some groups. Despite overall demographic similarity, we find differences in perceptions about these decisions in follow-up data.

4.2 Differences in Perception

Prior work and the intuition of the authors led to this study's focus on the cost/benefit analysis around these security advices. Regardless of which group's perceptions are more in line with reality, something that is mostly out of the view of this study, it is very likely that each group views the benefits, costs, and risks involved in the decision as different, which could at least in part be leading to the divergence in behavior.

Our first three hypotheses each focus on one of the three tenets of our study framework: benefit, cost, or risk. For all three, the guiding principle is that those who follow each advice are expected to have perceptions that are more supportive of adhering to the advice than the No groups. Please note that though only significant statistical results are detailed in this section due to space constraints. The results of all tests performed for this section can be found in the Appendix.

4.2.1 Benefits

As a core component of most rational decision models, it is natural to look at the benefits of a decision as perceived by those who are asked to make the decision. Specifically, for one to convince a person to do something, one must convince them that it is in their interests to do it. Through our design, we look at two kinds of benefits: the *Benefits of Following* (the security advice) and the *Benefits of Not Following* (the advice).

As explained in Section 3, each variable is defined using a single survey instrument that measures the variable on a 4-point Likert scale. Summaries of ratings for *Individual Benefit of Following* from each group, along with the results of a Mann-Whitney U-Test [18] comparing the response distributions of each Yes and No group can be seen in Table 3. Mann-Whitney U-Tests are appropriate for our data because the responses are independent and in the form of an ordinal scale. The test is non-parametric and measures if one distribution has a significantly higher median than the other, which would indicate, in our case, that one group rated the variable significantly higher than the other group. To analyze effect size, we use Cohen's d defined using the U-Test's Z score, divided by the square root of the number of samples compared [23].

	Yes	No	U-Test		
	Avg.(Med.)	Avg.(Med.)	U	Sig.	d
Upd.	3.77(4)	2.97(3)	274.5	<0.001	0.51
P.M.	3.78(4)	2.50(2.5)	154.5	<0.001	0.73
2FA	3.71(4)	2.90(3)	243.5	<0.001	0.49
Chg.P.	3.47(4)	2.53(3)	256	<0.001	0.57

Table 3: Rating summaries for *Individual Benefit of Following* for each group with U-Tests comparing the distribution between each Yes (those who follow the advice) and No (those who do not follow) groups. Effect size is measured with Cohen's *d*.

As can be seen in Table 3, for all advices, the Yes group rate their perceived benefit of following the advice as significantly higher than the No group, with most effects measuring "medium" (0.5) and one approaching "large" (0.8). This is unsurprising for the Yes groups since we expect that they are making a decision that they at least think benefits them. What's interesting here is the significantly lower ratings given by the No groups when asked to project the benefit they expected to receive from making the opposite decision of what they reported. As prior work has suggested [12, 13], these results support the idea that, at least in the eyes of some computer users, following security advice may just not be beneficial. This finding also supports our Hypothesis 1a, "For all decisions, the *Benefits of Following* will be seen as higher by the Yes groups compared to the No groups." Interestingly, ratings for *Social Benefit of Following* are not significantly different between groups for any advice, indicating that both see the benefits to "users of other computers" from each secure behavior as about the same. Of course, it could be that our samples are too small to show a significant effect and/or participants had a hard time conceptualizing the social benefits.

If one is interested in motivating more adherence to these and similar advices, this result suggests a gap between some users in how much benefit they see in adhering to these elements of advice. Addressing this gap through informational campaigns or other interventions may help, but providing

Advice	Group	n	Gender		Age		Comp. Expertise		Sec. Expertise		How Often Use Comp.	
			Male	*Female*	*Avg.*	*St.D.*	*Avg.*	*St.D.*	*Avg.*	*St.D.*	*Avg.*	*St.D.*
Update	Yes	39	20	19	38.4	14	4.15	0.7	3.56	0.8	4.79	0.4
	No	30	12	18	35.8	11	3.77	0.8	2.93	0.6	4.4	0.9
Password Manager	Yes	41	19	22	33.2	8.7	4.24	0.6	3.63	0.9	4.61	0.4
	No	38	16	22	34.0	9.7	4.3	0.7	3.50	0.7	4.79	0.4
2FA	Yes	36	20	16	36.6	13	4.31	0.7	3.86	0.9	4.69	0.5
	No	31	19	12	32.9	9	4.26	0.7	3.77	0.7	4.58	0.6
Change Passwords	Yes	37	20	17	36.0	10	4.22	0.6	3.78	0.8	4.73	0.6
	No	38	19	19	34.1	9.6	4.05	0.7	3.39	0.8	4.68	0.5

Table 2: Sample demographics for all groups used in this paper's analysis. "Comp[uter] Expertise", "Sec[urity] Expertise", and "How Often [Do You] Use [the] Comp[uter]?" are all measured on a 5-point Likert scale. The expertise questions are anchored from 1 = Very Poor to 5 = Very Good. The final question is anchored 1 = Never to 5 = All the Time, with the most common responses being "Often" or "All the Time."

better security tools, options, and education could go further towards increasing the adoption of secure behavior.

	Yes	No	U-Test		
	Avg.(Med.)	Avg.(Med.)	*U*	Sig.	*d*
Upd.	1.51(1)	2.13(2)	347.5	0.002	0.38
P.M.	1.68(1)	2.70(3)	302	<0.001	0.49
2FA	1.59(1.5)	2.62(3)	161.5	<0.001	0.61
Chg.P.	1.70(2)	3.03(3)	176	<0.001	0.66

Table 4: Rating summaries for *Individual Benefit of Not Following* for each group with U-Tests comparing the distribution between each Yes (those who follow the advice) and No (those who do not follow) groups. Effect size is measured with Cohen's *d*.

On a similar note, we do find significant differences between Yes and No groups' ratings on the variable *Individual Benefit of Not Following*, which supports our Hypothesis 1b. For updating, the effect is somewhat low, though still well above the "small" threshold (0.2), but other advice has solidly "medium" effects. Like before, there are no significant differences for *Social Benefits of Not Following*. As seen in Table 4, No groups consistently self-rate the benefits they receive from not following as significantly higher than the benefits the Yes groups' participants project they would receive from altering their behavior (i.e., to no longer following the advice). Like the ratings for *Individual Benefits of Following*, it should not be all too surprising that participants rate the benefits of their decision highly. If they thought the benefits were low, they likely would not be making the decision they claim they are. Still, for benefits, there is a perceptions gap when it comes to not following, as much as there is a perceptions gap for following. If those who do not behave securely see a lot of benefit in doing so, that must be addressed to alter their actions if so desired.

4.2.2 Risks

In addition to benefits, we also look at ratings of risk for more fine-grained insight into participants' considerations with respect to the tested advice. In the realm of security behavior, risk perception is a particularly important component to individuals' decisions as many behaviors are explicitly done to protect against a risk.

First, we analyze the perceptions of *Risks of Not Following*, covered by Hypothesis 2a. Table 5 shows the summaries for ratings to both *Individual Risk of Not Following* and *Social Risk of Not Following* along with U-Tests comparing each Yes groups' distribution with its corresponding No groups' distribution. Cohen's d is used to interpret effect size.

		Yes	No	U-Test		
		Avg.(Med.)	Avg.(Med.)	*U*	Sig.	*d*
Individ.	**Upd.**	3.42(4)	2.77(3)	336.5	0.002	0.37
	P.M.	2.88(3)	1.80(2)	302.5	<0.001	0.52
	2FA	3.42(3)	2.61(3)	243.5	<0.001	0.53
	Chg.P.	3.14(3)	2.63(3)	440.5	0.003	0.34
Social	**Upd.**	2.67(3)	1.76(1)	262.5	<0.001	0.44
	P.M.	1.92(2)	1.29(1)	409	0.002	0.37
	2FA	2.48(3)	1.79(2)	289	0.013	0.32
	Chg.P.	1.70(1)	1.29(1)	483	0.044	0.24

Table 5: Rating summaries for *Individ[ual] Risk of Not Following* and *Social Risk of Not Following* for each group with U-Tests comparing the distribution between each Yes (those who follow the advice) and No (those who do not follow) groups. Effect size is measured with Cohen's *d*.

Like with benefits, it is natural that those who follow each advice would see the risks of stopping that behavior as high since they are likely following to protect themselves from risks. In all cases, across both individual and social concerns, the Yes groups consistently rate the risks of no longer following the group's target advice as higher than the risks reported by those who already do not follow the target advice. This supports our hypothesis 2a, which states "For all decisions, the *Risks of Not Following* will be seen as higher by the Yes groups compared to the No groups." Effect sizes were sometimes low in these comparisons, but generally "medium."

As we stated before, we are not attempting to test the correctness of either group's perceptions, which would require data different than what was collected for this study. With that in mind, there is still much to learn from this result. There are many ways of interpreting the gap in risk perception between groups. On one hand, the risks could be low and those who follow the advices are exaggerating, as shown by the ratings from individuals who are *actually* at risk (No groups). In this view, one must assume that those who do not follow each advice are correctly experiencing the threat. This is where the alternative view comes in: it's possible, some may say, that those who do not follow have just not yet been affected, causing them to underestimate the risk of their behavior.

The existence of this perception gap calls for more research. In particular, as noted in prior work [12], identifying the

reality of risk faced by users is key to identifying which group is "correct." Knowing this can help further inform how (or where) to motivate behaviors that will increase security for users in a real way. Regardless, the risk perception gap can still be approached using the current state of knowledge. The authors also contend that these results could suggest that the targets of interventions (i.e., users who do not follow security advice) do not view their behavior as risky, despite the large amount of information and advice available online that should convince them otherwise. Thus, if the goal is altering these individual's decision, doing so may require new or alternative approaches. That said, the lower effect sizes compared to the other results presented up until this point could mean less of a gap here than for benefits.

It is also important to consider the perceived *Risks of Following* each advice since some may consider the tool or behavior risky. There are only strongly significant differences between groups on *Individual Risk of Following* for one advice: using a password managers ($U = 342.5$, $p < 0.001$. $d = 0.49$). The distribution of responses for the password manager Yes and No groups can be seen in Figure 1, which highlights the divergence in responses between the two groups. One other advice, changing passwords frequently shows a weaker, but significant difference for *Individual Risk of Following*($U = 498.5$, $p = 0.014$, $d = 0.28$). No other advice shows any significance in differences between groups on this variable. *Social Risk of Following* shows no significant differences for any advice.

Thus, our Hypothesis 2b is only partially supported by the data, particularly in the case of using password managers. Thinking about the function of a password manager in particular brings some insight. Password managers centralize passwords, an action some participants may view as risky, therefore increasing perceptions of risk of using the tool, especially among those who don't use it. Section 4.3.2 provides more information about possible reasons to explain this divergence.

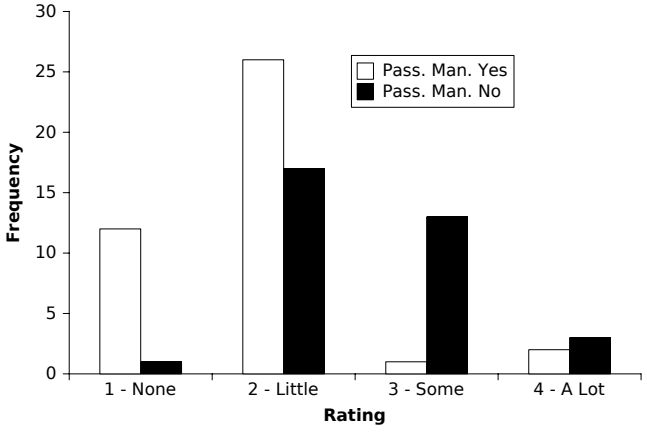

Figure 1: Response distributions representing the variable *Individual Risk of Following* for the password manager Yes (use) and No (don't use) groups.

4.2.3 Costs

Finally, besides benefits and risk, many decisions have some kind of cost associated with them. For example, updating one's system may take time in the form of a restart, or using 2FA on your phone may cost money in the form of charges for text messages. These costs will certainly play a role in the decision being made, thus we examine the cost ratings along with ratings of benefit and risk. Table 6 shows the summaries for the variable *Cost of Not Following* (both *Individual* and *Social*) along with U-Test results comparing the distributions of responses from the Yes group of each advice with the distribution of its corresponding No counterpart.

		Yes	No	U-Test		
		Avg.(Med.)	Avg.(Med.)	U	Sig.	d
Individ.	**Upd.**	2.95(3)	2.00(2)	247.5	<0.001	0.48
	P.M.	3.15(3)	1.75(1)	244.5	<0.001	0.60
	2FA	1.76(1)	1.57(1)	446.5	0.451	0.09
	Chg.P.	2.28(3)	1.61(1)	425.5	0.003	0.35
Social	**Upd.**	2.32(2)	1.59(1)	248	0.001	0.41
	P.M.	1.84(1)	1.03(1)	354	<0.001	0.49
	2FA	1.69(1)	1.41(1)	343	0.356	0.12
	Chg.P.	1.50(1)	1.24(1)	525.5	0.174	0.16

Table 6: Rating summaries for *Individ[ual] Cost of Not Following* and *Social Cost of Not Following* for each group with U-Tests comparing the distribution between each Yes (those who follow the advice) and No (those who do not follow) groups. Effect size is measured with Cohen's *d*.

As can be seen in Table 6, some advice has significant differences in how participants in each group rate the costs of not following the advice. Updating and using a password manager shows the strongest differences, with there being divergence on both the individual and social forms of the variable. Effect sizes are "medium" in each case. Changing passwords also shows significant differences, but only for the *Individual Cost of Not Following*. Here the effect size is smaller than for other differences in cost. For updating, many users may see a benefit in terms of performance when updating and so see not updating as incurring them a cost (i.e., in performance). Similarly, password managers help with things like account creation and log in, so they provide a convenience benefit in addition to security benefit. Thus, it is likely that those who stopped using a password manager would feel a cost in terms of time and/or effort. What's interesting is that there are differences between the groups on costs for some of the advice, which could mean that there is a real benefit incurred by updating and/or using a password manager that is not known until trying. Overall, these results somewhat support our Hypothesis 3a, just not for all cases as predicted.

There is only one strongly significant result when comparing ratings from the Yes and No groups for the variable *Cost of Following*, which is *Individual Cost of Following* for changing passwords. The No group rates this significantly higher than the Yes group, with averages of 2.97 and 2.35 respectively (U=449.5, p=0.005, d=0.33). Two other elements of advice also have weaker, but signifigant differences on this variable: using a password manage (U=533, p=0.011, d=0.28) and using 2FA (U=405.5, p=0.036, d=0.26). For the individual cost of updating and social phrasings of the *Cost of Following* for all pieces of advice, differences are not significant. Thus, we only have limited data to directly support the hypothesis "For all decisions, the *Costs of Following* will be seen as higher by the No groups compared to the Yes

groups."

Though the ratings from participants provide a window into their minds, the quantitative data is limited in richness due to its nature. As such, we supplement our numerical data with open-ended responses as to *why* participants made the decision they did, as explained in Section 3. Analysis of these comments helps answer some of the questions presented by quantitative analysis.

4.3 Why Do They Do What They Do?

In addition to finding perception differences, this study hopes to shed some light on the reasons people have for their decisions, which can help us explain some of the gaps. Hypotheses 4a and 4b deal with this aspect of the study and are supported using analysis of the qualitative data. Responses are coded using the process described in Section 3.2. To examine how the reasons provided by each group differ, and how well our hypotheses predict our results, in this section, we present the most prominent and noteworthy codes assigned to comments from each group.

4.3.1 Updating

Keeping one's software up to date is one of the most commonly recommended practices from security experts to keep data and machines protected. When comparing the reasons for each decision (i.e., to update or not), we find stark differences, but also some interesting similarities.

First, a large number of those who update said they do so for security purposes. Approximately 49% of all 39 comments received from Yes group participants mention increasing some kind of security. Additionally, good performance, specifically avoiding bugs and software issues is a chief concern for the group of participants who update. Twenty-two of 39 comments mention avoiding bugs and/or issues, making up 56% of the comments from this group. Ten (26%) comments mention wanting to get the most recent changes, while 7 (18%) indicate a desire to avoid malware specifically.

Unsurprisingly, these codes were not assigned to any comments from the No group. Instead, common concerns for that group are getting a convenience and/or avoiding an inconvenience, not finding a need [to update], or not being willing to put in the effort involved. Five of 30 comments (15%) mention not needing to update, while another five comments mention being too lazy to update and/or updates being too much work to apply. 23% of the comments allude to or mention avoiding an inconvenience and/or getting a convenience by not applying updates. Interestingly, 13% of the comments from those who do update also bring up avoiding an inconvenience/getting a convenience. It would seem that both groups see some convenience in their decision, be it through avoiding undue effort, as in the case for the No group, or through getting the latest features, as for the Yes group.

Those who do not update have many other specific reasons for their decision. Avoiding harm (3 comments), avoiding change (5 comments), and finding updates too frequent (4 comments) are also common reasons from the No group, showing the spread of concern among these individuals. By contrast, most of those who update report similar reasons (i.e., security, best features, avoid software faults) for their decision.

Looking at updating, both our hypotheses for this aspect of the study hold up. Hypothesis 4a states "Those who follow each advice will do so, generally, to increase their security and/or for convenience purposes," which is supported by the large number of comments from the Yes group saying they update to increase their security or to avoid software issues. Of those who do not update, many choose that route to get a convenience/avoid an inconvenience, supporting Hypothesis 4b. Many others also mention a confidence in their current approach by saying or suggesting they have no need to update. It should be noted that 3 comments bring up a specific bad update in the past as a reason for their update avoidance, so it's possible some participants' skepticism is warranted or, at least understandable from a rational decision standpoint, as suggested in prior work [29, 28].

4.3.2 Using a Password Manager

Password managers help create and manage passwords for online accounts by allowing automatic form filling, which alleviates the need for users to remember many, long, complex (and therefore secure) passwords. "Secure" password managers help increase overall privacy by affording users the ability to auto-generate and auto-fill hard-to-crack passwords on all their accounts. Recommended password managers encrypt the stored data to reduce the obvious security risk introduced by storing all passwords in a single, noticeable, predictable place. Password managers that do not encrypt passwords are generally considered insecure, but our study specifically asks participants if they use more secure password managers (e.g., LastPass).

Those who use a password manager report the convenience added by the tool (i.e., automatic form-filling) as a reason for using in an overwhelming majority of their comments. Thirty-seven of 40 comments (93%) from those who use a password manager mention the added convenience of the software. 55% of comments from the same group indicate the added security they get from using their password manager as a reason for their decision to use.

By contrast, 45% of those who do not use a password manager say they avoid them to avoid a security risk, showing that many in the No group feel that password managers are not worth the added benefit at log-in because they think the tool opens them up to attack. Twelve (32%) of 38 comments from the No group specifically mention avoiding centralizing their passwords as a reason not to use a password manager. Calling back to prior results from this study, these comments can shed some light on the significantly higher ratings for the *Individual Risk of Following* from those who do not use a password manager compared to those who do use one. It seems that a large proportion of those who do not use a password manager explicitly do not because they view the tool as a security risk.

Additionally, half of the comments mentions a confidence in the participant's current security/password mechanism. These approaches include remembering passwords (which could lead to insecure passwords used on websites due to cognitive limitations for individuals to remember log-ins) and writing passwords down in a "secure" place, which may seem satisfactory, but ignores risks from local threats and could also lead to bad passwords due to complacency.

Thinking to our Hypotheses 4a and 4b, these results sup-

port those predictions. Users of password managers report becoming users in a majority of cases examined because of the security and convenience they feel they get from their action, thus supporting 4a. On the other side, those who do not use a password manager in many cases do so because they feel their current method of password management is sufficient (i.e., confidence in current behavior), partially supporting 4b. That all said, password managers were different from the other advices we tested in that many non-users reported their impression of password managers as a fundamental risk as a reason for not using them. This is reflected in the qualitative data presented in Section 4.2.2.

4.3.3 Using 2FA

Two-factor authentication (2FA) is another common technique for increasing account security. In addition to a username and password, users of 2FA are sent a one-time password through email, SMS, etc. that is used in the specific instance of that log in. The addition of the one-time password, which is only good for the single log in attempt, increases security by adding another factor (of authentication) that must be stolen by a would-be attacker. If a hacker, for example, gets access to your user-name and password, by using 2FA, they would also need access to the account and/or device you use to receive your one-time passwords to be able to access your account.

A large proportion of the 36 comments from participants in the group who report using 2FA say they do so to increase their security (86%) and/or because it's safer than the alternatives they're aware of (61%). Additionally, 25% say they use 2FA because it "feels better" than not using 2FA. Overall, these comments suggest that 2FA users are strongly motivated by the security benefits they see in the technique. This should not be surprising as 2FA is less commonly used and is known for increasing security, so those who *do* use it are likely to be drawn by that prominent benefit.

On the flip side, 48% of the 31 comments from the 2FA No group say they do not use 2FA to avoid an inconvenience and 23% mention avoiding a cost. In both cases, the most common cost and/or inconvenience is the need for a second factor, which slows log-in. Additionally, 26% of the comments mention that the participants' current approach is good enough, 19% say they do not see the risks of not using 2FA and/or don't care if they're hacked, and 13% allude to or say there is no need for using 2FA.

Like before, these findings broadly support Hypotheses 4a and 4b. The Yes group for 2FA greatly values the security they get from using 2FA, but unlike updating and using a password manager, none think 2FA offers them a convenience. Convenience or more specifically the avoidance of the inconvenience of 2FA is a chief concern among those who don't use 2FA. Not seeing a need to use 2FA and the idea that their current approach is good enough (compared to 2FA) also influence the No group.

4.3.4 Changing Passwords Frequently

Frequently changing passwords, though not a common advice from experts, is seen as a secure behavior in the eyes of many users [15], likely due to password changes being recommended in corporate environments and/or after a security breach. Changing passwords frequently is not likely to help protect an individual account, assuming all passwords used are of sufficient security. The security benefits come in when the attacker may have access to your *current* password, but by changing it, you thwart their attack.

Like the use of 2FA, those who frequently change their passwords commonly cite the added security they get from doing so, as was the case for 26 (72%) of 36 comments from the Yes group. 19% of the comments from this group specifically mention increased account security, and 22% mention avoiding theft and/or unauthorized access of their account. None mention a convenience increase as a reason for their decision to use.

For those who do not change their passwords frequently, also like 2FA, many (53% of 38 comments) say they do so to avoid an inconvenience. Other concerns, like confidence in their current approach (13%) and seeing a low risk of attack (18%) are also common reasons for not changing passwords. Unlike 2FA, though, many (39%) comments say they do not change passwords often because doing so is too hard to remember and/or their passwords would be too hard to remember if they did. Also, interestingly, 32% of comments from this group mention not having problems before as a reason not to start changing passwords (and therefore continue **not** changing passwords), while only 6% of those who don't use 2FA mentioned such a theme in their comment. It could be that, due to changing passwords being less "work" than using 2FA, participants who do not follow the advice feel more reason to justify their decision in another way, in this case using an argument of "if it ain't broke, don't fix it," while those who do not use 2FA feel justified in avoiding the somewhat substantial extra cost of enabling the feature.

The Yes group's focus on the perceived security benefits of changing passwords frequently supports, at least in part Hypothesis 4a. By worrying about the inconvenience of changing passwords frequently and not seeing much risk in their behavior, the comments from the No group also supports Hypothesis 4b.

4.3.5 Social Content in Comments

One very strong theme across all comments is the focus on the individual in the reasons given. Only 13 of 290 (4.5%) of all comments mention some social motivation behind the decision, all from Yes groups. This is in line with prior work showing the positive effect of social motivation around computer security [6], since the few comments that did mention a social motivation were all from participants that followed security advice. Examples of social motivations in comments include the desire to protect family/other users (5 comments), trust in developers (2 comments), acting on a friend's/family member's recommendation (4 comments), and concern for their place in the Internet/network in general (2 comments). With this lead and hints from prior work, we further investigate the individual/social motivation divide.

4.4 Individual vs. Social Concerns

As described in the Methods section, each component of our decision model is toned in both an individual and social context. All participants were asked for an individual and social rating for each component (i.e., benefit/cost/risk) related to following and not following the advice. Figure 2 shows the average overall rating (across all groups) for each variable in our study plotted together to contrast the difference between

averages for individual and social phrasings.

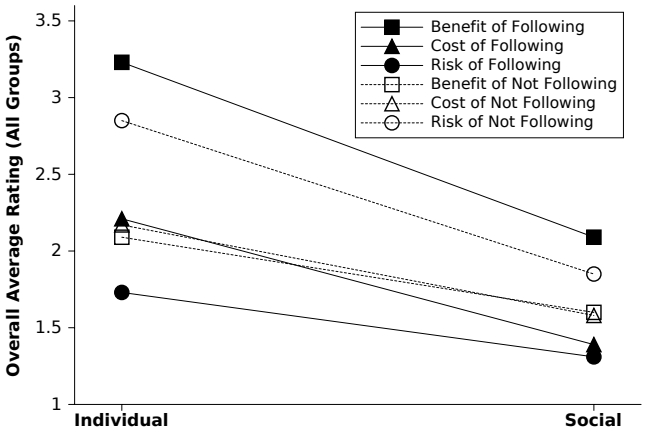

Figure 2: Plot of average overall ratings for each variable, arranged to show the consistently lower social ratings compared to individual ratings. A sign test of each variable pair (ind. vs. soc.) found significant ($p < 0.001$) differences for all variables.

As can be seen in Figure 2, for each variable, the individual phrasing average is higher than the corresponding social phrasing's average. To statistically test these differences, we use a sign test [7]. Put simply, the sign test determines if one variable from the pair tested is rated consistently higher than the other. A low p value for the sign test indicates that participants in the sample consistently rated one variable from each pair as higher than the other variable in the pair. We meet the assumptions for this test since our data is ordinal and observations from different participants are independent, thus the differences in their individual and social scores are also independent.

In the context of our study, we use the sign tests to determine if the differences demonstrated by the averages plotted in Figure 2 are representative of statistically significant and consistent differences between individual and social ratings, regardless of advice, aspect of decision, or context (i.e., following vs. not following the advice). Data was aggregated for each variable across all 8 groups, and the sign test compares *Individual* with *Social* ratings. For all pairs tested, we find strongly significant differences ($\forall$, $Z <$ -5, $p < 0.001$), indicating that ratings for individually phrased variables are consistently higher than the socially phrased version's ratings. Effect sizes measured using Cohen's d were greater than 0.5 for all tests except *Benefit of Not Following*, which was 0.36, indicating that the differences between groups could be considered "medium." Full results of these tests can be found in the Appendix.

Lower social ratings than individual indicate that most participants may give more consideration to how the option of following each tested security advice affects them than how it affects others. As prior work has indicated, social motivation are stronger regulators of behavior than individual motivations [27, 20]. Computer security is ripe for social considerations as one's security behavior can have an effect on other's security, especially if your behavior causes a breach of some kind. For example, if by not updating your operating system, your machine is infected and becomes a member of a malicious botnet, your decisions will have affected others when the botnet is used against websites or other web-services utilized by other computer users. Thus, increasing the strength of social considerations around computer security is not only possible, but preferable to focusing on individual considerations when trying to motivate good security behavior.

Though there are strong differences when aggregating, we also use sign tests to compare the *Individual* and *Social* ratings for each variable separated by group to see how the overall results hold up when looking at specific contexts. In most cases, the difference between individual and social ratings holds in significance. Only 23 of 48 tests have significant values greater than 0.001. Sixteen of those tests show weaker, but nonetheless significant differences, including 11 tests resulting in $p \leq 0.008$ and 5 additional tests returning $p < 0.04$. The remaining 7 cases do not show significant differences, but these are mostly *Benefit* and/or *Cost of Not Following*, which could be hard concepts for some participants to wrap their heads around. Additionally, larger samples may show stronger differences for these variables. These findings support our final hypothesis, "Social considerations will be lower than individual concerns for all decisions." As before, the full results of these tests can be found in the Appendix.

5. DISCUSSION AND LIMITATIONS

Overall, our findings enlighten in solving the problem of motivating secure behavior. Even as the science of security improves and new, better tools are released, the task of getting users to take up these tools and techniques will always exist. Discovering trends in perceptions around these decisions and more importantly using them to help develop strategies of persuasion are both key towards a more secure ecosystem.

Central is the propensity of each group to see the benefits of their decision as higher than their counterparts predict their benefits would be. Though unsurprising, it is important for those giving security advice to keep in mind that even though you, as an adherent to a behavior see certain benefits, others, particularly those who do not adhere are likely to not see the same benefits. Though this may suggest a simple solution is to better inform users about benefits (which assumes the No groups are wrong in their perceptions), prior work argues that such an approach is likely to fail [12, 13, 5], indicating that the users (those who do not follow the advices included) are at least aware of the benefits involved and do not need to be simply informed. Besides simple ignorance, there may be many other reasons for these perception gaps. It could be that some do not realize the value of the benefits, or the benefits are actually not as high as the Yes groups seem to think. These and other explanations for the differences in benefits require a different solution than simply disseminating information. Instead, the task calls for a nuanced, issue-tailored approach that addresses what users are likely thinking and what they actually experience to help them overcome the barriers to desired behavior.

Risk perception is important too, as to be expected in the realm of security decision-making. Like with benefits, we found that Yes and No groups felt differently about the risks they were protected from by following each behavior. It is very hard to know which group is more accurately estimating the risks involved as there is limited data on the

costs and risks experienced by an average, individual user. Though there is much data on the macro-level (e.g., number of attacks per year, accounts compromised every day, etc.), there has been no large-scale, regular data collection to give empirical and scientific power to statements about the danger of general online security risks to a particular user. As researchers, we try our best to estimate these risks, but without hard, consistent data, any advice we give is on some level speculative and based on our incomplete picture of what users face. Calls for this kind of data are not new, but have thus far gone unanswered.

The convergence of most participants' justifications for their decisions around the topics of security and/or (in)convenience is also notable as the convenience/security trade-off is a commonly discussed concept around computer security [26, 30, 21]. In general, many note that security requires some kind of inconvenience while taking a more convenient route will likely prove less secure. For example, it is much easier to make and manage a single account for a shared machine, but such a set-up makes activity and data from different users visible to others, resulting in less security than individual accounts. In some cases of our study, such as changing passwords frequently and using 2FA, most who followed the advice say they do so for a security benefit, while most who do not follow say their decision is to avoid an inconvenience, suggesting participants making these decisions are considering a security/convenience trade-off. Time was a very common theme, with participants citing a lack of time to follow the tested advice. As one non-updating participant put it: "I'm busy, dang it!"

Many No group comments express similar sentiments. Use of a password manager also plays into this paradigm, but shuffles it due to the specific functionality of password managers. Many of those who do not use password managers report avoiding the security risk of centralization as their concern with the tool, while many users cite the convenience benefit afforded by auto-login features. Updating is also reported to come with benefits (e.g., in the form of better software performance) that are appreciated by participants. These findings suggest that motivating more secure behavior could be done with better management of the convenience/security trade-off considerations being made for particular context.

Finally, our results show that individual rather than social concerns are rated higher in quantitative data and are more prominent in the qualitative data. Though the lack of social comments could be due to question wording (i.e., the open-ended question's phrasing may encourage responses biased towards individual concerns), the existence of several comments that **do** mention a social motivation and the quantitative results related to social vs. individual concerns both show that many participants are thinking predominately about themselves when making these decisions.

Psychology has long studied the occurrence of prosocial behavior [20], in no small part because such behavior is very beneficial to society as a whole and so society is inclined to encourage it in individuals where possible. Newer research has pointed to the power of social motivations [27]. If the social consciousness of these decisions could be increased, it is likely that some users will be motivated to follow despite the costs they may incur. Like before, more data on the real risks and ramifications of security threats and efficacy of various behaviors in protecting adherents is important here because knowing the social effects is key to properly adjusting user's perceptions, when necessary.

Our approach is not without its limitations. Though we were able to find statistically significant differences in many places, additional data could generate new findings or provide insight in existing results. In particular, larger and more varied samples could garner larger effect sizes than those reported in this study, which were generally "medium." In addition, examination of more advices and contexts (e.g., perceptions of benefits/risks/costs for specific kinds of devices) could also expand the picture. An expanded decision-making framework may provide more insight, but would likely require a larger study from the design presented and used here, introducing different limitations. Finally, as Mechanical Turk's user-base may not be representative of the general population, replication of this study with more samples would help generalize the findings.

6. CONCLUSION

Our results show differences in the perceptions of benefits, risks, and costs associated with decisions to adhere to a variety of security behaviors that are commonly recommended by experts. Both those who do and do not follow each advice report that their current decision gets them more benefit than if they changed. Those who follow rate the risks of changing their decision as much higher than the risks reported by those who do not follow. Costs of not following are also seen as higher by most that follow compared to those who do not. When looking into the reasons participants gave for their decisions, we find strong trends highlighting the convenience/security trade-off. The value of convenience in particular may be used to help motivate the use of security tools and techniques. Finally, we found that individual concerns are rated consistently higher than social concerns. Increasing social motivations could motivate more secure decision-making, according to theory from prior work [27].

Additional data regarding the real benefits/risks/costs of these and related contexts, not just perceptions of them are needed to help better paint the complete picture of what is happening in users' minds and address the gaps identified. Nonetheless, this study has provided insight into user motivation to guide future efforts towards the broader goals of usable security.

7. ACKNOWLEDGMENTS

The authors thank the reviewers and the paper's shepherd, Lujo Bauer for their helpful guidance with this paper. This work is supported by the National Science Foundation under Grant no. CNS-1343766 and by GAAAN Fellowship no. P200A130153. Any opinions, findings, or recommendations expressed are those of the authors and do not necessarily reflect the views of the funding agencies.

8. REFERENCES

[1] A. Adams and M. A. Sasse. Users are not the enemy. *Communications of the ACM*, 42(12):40–46, 1999.

[2] R. A. Armstrong and A. C. Hilton. Nonparametric correlation coefficients. *Statistical Analysis in Microbiology: Statnotes*, pages 91–94.

[3] F. Asgharpour, D. Liu, and L. J. Camp. Mental models of security risks. In *Financial Cryptography*

and Data Security, pages 367–377. Springer, 2007.

[4] L. J. Camp. Mental models of privacy and security. *Available at SSRN 922735*, 2006.

[5] N. Clarke, S. Furnell, G. Stewart, and D. Lacey. Death by a thousand facts: Criticising the technocratic approach to information security awareness. *Information Management & Computer Security*, 20(1):29–38, 2012.

[6] S. Das, T. H.-J. Kim, L. A. Dabbish, and J. I. Hong. The effect of social influence on security sensitivity. In *SOUPS*, pages 143–157, 2014.

[7] W. J. Dixon and A. M. Mood. The statistical sign test. *Journal of the American Statistical Association*, 41(236):557–566, 1946.

[8] D. Florencio and C. Herley. A large-scale study of web password habits. In *Proceedings of the 16th international conference on World Wide Web*, pages 657–666. ACM, 2007.

[9] M. Harbach, S. Fahl, and M. Smith. Who's afraid of which bad wolf? a survey of it security risk awareness. In *Computer Security Foundations Symposium (CSF), 2014 IEEE 27th*, pages 97–110. IEEE, 2014.

[10] M. Harbach, M. Hettig, S. Weber, and M. Smith. Using personal examples to improve risk communication for security and privacy decisions. In *Proceedings of the 32Nd Annual ACM Conference on Human Factors in Computing Systems*, CHI '14, pages 2647–2656, New York, NY, USA, 2014. ACM.

[11] M. Harbach, E. von Zezschwitz, A. Fichter, A. De Luca, and M. Smith. It's a hard lock life: A field study of smartphone (un)locking behavior and risk perception. In *Symposium on Usable Privacy and Security (SOUPS 2014)*, pages 213–230, 2014.

[12] C. Herley. So long, and no thanks for the externalities: the rational rejection of security advice by users. In *Proceedings of the 2009 workshop on New security paradigms workshop*, pages 133–144. ACM, 2009.

[13] C. Herley. More is not the answer. *IEEE Security & Privacy*, (1):14–19, 2014.

[14] A. E. Howe, I. Ray, M. Roberts, M. Urbanska, and Z. Byrne. The psychology of security for the home computer user. In *Security and Privacy (SP), 2012 IEEE Symposium on*, pages 209–223. IEEE, 2012.

[15] I. Ion, R. Reeder, and S. Consolvo. "... no one can hack my mind": Comparing expert and non-expert security practices. In *Eleventh Symposium On Usable Privacy and Security (SOUPS 2015)*, pages 327–346, 2015.

[16] R. Kang, L. Dabbish, N. Fruchter, and S. Kiesler. "my data just goes everywhere:" user mental models of the internet and implications for privacy and security. In *Eleventh Symposium On Usable Privacy and Security (SOUPS 2015)*, pages 39–52, 2015.

[17] M. Madden and L. Rainie. Americans' attitudes about privacy, security and surveillance. Online at: `http://www.pewinternet.org/`, May 2015.

[18] H. B. Mann and D. R. Whitney. On a test of whether one of two random variables is stochastically larger than the other. *The annals of mathematical statistics*, pages 50–60, 1947.

[19] P. E. McKight and J. Najab. Kruskal-wallis test. *Corsini Encyclopedia of Psychology*, 2010.

[20] L. A. Penner, J. F. Dovidio, J. A. Piliavin, and D. A. Schroeder. Prosocial behavior: Multilevel perspectives. *Annu. Rev. Psychol.*, 56:365–392, 2005.

[21] S. Prabhakar, S. Pankanti, and A. K. Jain. Biometric recognition: Security and privacy concerns. *IEEE Security & Privacy*, (2):33–42, 2003.

[22] L. Rainie, S. Kielsler, R. Kang, and M. Madden. Anonymity, privacy, and security online. Online at: `http://www.pewinternet.org/`, September 2013.

[23] R. Rosenthal, H. Cooper, and L. Hedges. Parametric measures of effect size. *The handbook of research synthesis*, pages 231–244, 1994.

[24] B. Rossler and R. D. Glasgow. *The value of privacy*. Polity, 2005.

[25] A. Strauss and J. Corbin. Grounded theory methodology. *Handbook of qualitative research*, pages 273–285, 1994.

[26] L. Tam, M. Glassman, and M. Vandenwauver. The psychology of password management: a tradeoff between security and convenience. *Behaviour & Information Technology*, 29(3):233–244, 2010.

[27] T. R. Tyler. *Why people cooperate: The role of social motivations*. Princeton University Press, 2010.

[28] K. E. Vaniea, E. Rader, and R. Wash. Betrayed by updates: How negative experiences affect future security. In *Proceedings of the 32Nd Annual ACM Conference on Human Factors in Computing Systems*, CHI '14, pages 2671–2674, New York, NY, USA, 2014. ACM.

[29] R. Wash, E. Rader, K. Vaniea, and M. Rizor. Out of the loop: How automated software updates cause unintended security consequences. pages 89–104. USENIX Association, 2014.

[30] C. S. Weir, G. Douglas, M. Carruthers, and M. Jack. User perceptions of security, convenience and usability for ebanking authentication tokens. *Computers & Security*, 28(1):47–62, 2009.

APPENDIX

A. SURVEY INSTRUMENTS

Templates for all instruments used in this study are listed in the subsections below.

A.1 Intial Survey

The questions derived from thew following template were shown to 805 participants who initially enrolled in the study from Mechanical Turk.

1. What is your age?
2. What is your gender?
   - Male
   - Female
   - Other
3. Do you use a laptop or desktop computer that you or your family owns (i.e., not provided by school or work)?
   - Yes
   - No
4. How would you rate your general computer expertise?
   - Very Poor
   - Poor
   - Fair

_ Good
_ Very Good

5. How would you rate your computer security expertise?
   _ Very Poor
   _ Poor
   _ Fair
   _ Good
   _ Very Good
6. How often would you say you use the computer?
   _ Never
   _ Rarely
   _ Sometimes
   _ Often
   _ All the Time
7. Do you keep your computer's software up to date?
   _ Yes
   _ No
   _ I Don't Know
8. Do you use two-factor authentication (e.g., 2-Step Verification) for at least one of your online accounts?
   _ Yes
   _ No
   _ I Don't Know
9. Do you use a password manager (e.g., LastPass, OnePass, KeePass) to manage your online account passwords?
   _ Yes
   _ No
   _ I Don't Know
10. Do you change your passwords frequently?
    _ Yes
    _ No
    _ I Don't Know

A.2 Follow-Up Surveys

After groups were formed, the following templates were used to create surveys for each advice Yes and No group. To form each survey, replace [follow(ing) the advice] in the templates with each of the following phrases for the corresponding advice:

- **Update** - "keep(ing) your computer's software up to date"
- **Pass. Man.** - "us(e/ing) a password manager"
- **2FA** - "us(e/ing) two-factor authentication"
- **Change Pass.** - "chang(e/ing) your passwords frequently"

A.2.1 "Yes" Group Template

1. Please explain in a few sentences why you choose to [follow the advice].
2. How much would you say you are benefited by you [following the advice]?
   _ None
   _ Little
   _ Some
   _ A Lot
   _ Not Sure
3. How much would you users of other computers are benefited by you [following the advice]?
   _ None
   _ Little
   _ Some
   _ A Lot
   _ Not Sure
4. How much would you say you are cost or inconvenienced by you [following the advice]?
   _ None
   _ Little
   _ Some
   _ A Lot
   _ Not Sure
5. How much would you users of other computers are cost or inconvenienced by you [following the advice]?
   _ None
   _ Little
   _ Some
   _ A Lot
   _ Not Sure
6. How much would you say you are put at risk by you [following the advice]?
   _ None
   _ Little
   _ Some
   _ A Lot
   _ Not Sure
7. How much would you users of other computers are put at risk by you [following the advice]?
   _ None
   _ Little
   _ Some
   _ A Lot
   _ Not Sure
8. How much would you say you would be benefited if you did not [follow the advice]?
   _ None
   _ Little
   _ Some
   _ A Lot
   _ Not Sure
9. How much would you say users of other computers would be benefited if you did not [follow the advice]?
   _ None
   _ Little
   _ Some
   _ A Lot
   _ Not Sure
10. How much would you say you would be cost or inconvenienced if you did not [follow the advice]?
    _ None
    _ Little
    _ Some
    _ A Lot
    _ Not Sure
11. How much would you say users of other computers would be cost or inconvenienced if you did not [follow the advice]?
    _ None
    _ Little
    _ Some
    _ A Lot
    _ Not Sure

12. How much would you say you would be put at risk if you did not [follow the advice]?
    - None
    - Little
    - Some
    - A Lot
    - Not Sure
13. How much would you say users of other computers would be out at risk if you did not [follow the advice]?
    - None
    - Little
    - Some
    - A Lot
    - Not Sure

A.2.2 "No" Group Template

1. Please explain in a few sentences why you choose not to [follow the advice].
2. How much would you say you are benefited by you [following the advice]?
    - None
    - Little
    - Some
    - A Lot
    - Not Sure
3. How much would you users of other computers are benefited by you not [following the advice]?
    - None
    - Little
    - Some
    - A Lot
    - Not Sure
4. How much would you say you are cost or inconvenienced by you not [following the advice]?
    - None
    - Little
    - Some
    - A Lot
    - Not Sure
5. How much would you users of other computers are cost or inconvenienced by you not [following the advice]?
    - None
    - Little
    - Some
    - A Lot
    - Not Sure
6. How much would you say you are put at risk by you not [following the advice]?
    - None
    - Little
    - Some
    - A Lot
    - Not Sure
7. How much would you users of other computers are put at risk by you not [following the advice]?
    - None
    - Little
    - Some
    - A Lot
    - Not Sure
8. How much would you say you would be benefited if you did [follow the advice]?
    - None
    - Little
    - Some
    - A Lot
    - Not Sure
9. How much would you say users of other computers would be benefited if you did [follow the advice]?
    - None
    - Little
    - Some
    - A Lot
    - Not Sure
10. How much would you say you would be cost or inconvenienced if you did [follow the advice]?
    - None
    - Little
    - Some
    - A Lot
    - Not Sure
11. How much would you say users of other computers would be cost or inconvenienced if you did [follow the advice]?
    - None
    - Little
    - Some
    - A Lot
    - Not Sure
12. How much would you say you would be put at risk if you did [follow the advice]?
    - None
    - Little
    - Some
    - A Lot
    - Not Sure
13. How much would you say users of other computers would be out at risk if you did [follow the advice]?
    - None
    - Little
    - Some
    - A Lot
    - Not Sure

B. STATISTICAL RESULTS

This section contains statistics generated and tests performed for this study, including those not included in the paper's main text. Tables 7- 9 on the following pages contain the details for Mann-Whiteny U-Tests and sign tests used in this paper's analysis.

			. . . of Following					. . . of Not Following				
			Yes	*No*	U-Test			*Yes*	*No*	U-Test		
			Avg.(Med.)	Avg.(Med.)	*U*	Sig.	*d*	Avg.(Med.)	Avg.(Med.)	*U*	Sig.	*d*
Benefit	*Individ.*	Upd.	3.77(4)	2.97(3)	274.5	<0.001	0.51	1.51(1)	2.13(2)	347.5	0.002	0.38
		P.M.	3.78(4)	2.50(2.5)	154.5	<0.001	0.73	1.68(1)	2.70(3)	302	<0.001	0.49
		2FA	3.71(4)	2.90(3)	243.5	<0.001	0.49	1.59(1.5)	2.62(3)	161.5	<0.001	0.61
		Chg.P.	3.47(4)	2.53(3)	256	<0.001	0.57	1.70(2)	3.03(3)	176	<0.001	0.66
	Social	Upd.	2.71(3)	2.39(3)	338	0.286	0.14	1.40(1)	1.58(1)	371	0.371	0.12
		P.M.	2.08(2)	1.70(1)	498.5	0.155	0.17	1.39(1)	1.68(1)	511	0.142	0.18
		2FA	2.48(2)	2.29(2)	390	0.489	0.09	1.59(1)	1.92(1.5)	313.5	0.237	0.16
		Chg.P.	1.73(1)	1.48(1)	463.5	0.235	0.15	1.74(1)	1.58(1)	511	0.467	0.09
Risk	*Individ.*	Upd.	1.56(2)	1.72(2)	496.5	0.335	0.12	3.42(4)	2.77(3)	336.5	0.002	0.37
		P.M.	1.83(2)	2.53(2)	342.5	<0.001	0.49	2.88(3)	1.80(2)	302.5	<0.001	0.52
		2FA	1.56(1)	1.62(1)	498.5	0.729	0.04	3.42(3)	2.61(3)	243.5	<0.001	0.53
		Chg.P.	1.35(1)	1.71(2)	498.5	0.014	0.28	3.14(3)	2.63(3)	440.5	0.003	0.34
	Social	Upd.	1.13(1)	1.38(1)	369.5	0.047	0.25	2.67(3)	1.76(1)	262.5	<0.001	0.44
		P.M.	1.41(1)	1.53(1)	628	0.707	0.04	1.92(2)	1.29(1)	409	0.002	0.37
		2FA	1.31(1)	1.48(1)	433.5	0.47	0.09	2.48(3)	1.79(2)	289	0.013	0.32
		Chg.P.	1.19(1)	1.17(1)	628.5	0.709	0.04	1.70(1)	1.29(1)	483	0.044	0.24
Cost	*Individ.*	Upd.	2.03(2)	2.1(2)	527.5	0.444	0.09	2.95(3)	2.00(2)	247.5	<0.001	0.48
		P.M.	1.73(2)	2.18(2)	533	0.011	0.28	3.15(3)	1.75(1)	244.5	<0.001	0.60
		2FA	2.00(2)	2.39(2)	405.5	0.036	0.26	1.76(1)	1.57(1)	446.5	0.451	0.09
		Chg.P.	2.35(2)	2.97(3)	449.5	0.005	0.33	2.28(3)	1.61(1)	425.5	0.003	0.35
	Social	Upd.	1.22(1)	1.29(1)	431	0.781	0.04	2.32(2)	1.59(1)	248	0.001	0.41
		P.M.	1.28(1)	1.52(1)	565.5	0.213	0.15	1.84(1)	1.03(1)	354	<0.001	0.49
		2FA	1.52(1)	1.44(1)	403.5	0.786	0.04	1.69(1)	1.41(1)	343	0.356	0.12
		Chg.P.	1.28(1)	1.65(1)	491	0.073	0.21	1.50(1)	1.24(1)	525.5	0.174	0.16

Table 7: Rating summaries for all variables with U-Tests comparing the distribution between each Yes (those who follow the advice) and No (those who do not follow) groups. Effect size is measured with Cohen's *d*.

	. . . of Following						. . . of Not Following					
	Ind.>	Soc.>	Tie	*Z*	Sig.	*d*	Ind.>	Soc.>	Tie	*Z*	Sig.	*d*
Benefit	176	10	62	-12.10	<0.001	0.77	108	38	99	-5.71	<0.001	0.36
Risk	112	8	148	-9.40	<0.001	0.57	174	6	85	-12.45	<0.001	0.76
Cost	165	21	75	-10.49	<0.001	0.65	102	11	140	-8.47	<0.001	0.53

Table 8: Sign test results comparing *Individual* and *Social* ratings for each variable from all participants aggregated across both groups and all advice tested. Along with the *Z* and *p* values, we also show difference frequences to show how often participants' *Individual* ratings were higher, lower, or tied with their *Social* rating. Effect size is measured with Cohen's *d*.

			...of Following					...of Not Following				
			Ind.>	Soc.>	Tie	*Z*	Sig.	Ind.>	Soc.>	Tie	*Z*	Sig.
Yes Groups	Update	**Benefit**	25	0	10	-	<0.001	8	2	25	-	0.109
		Risk	17	1	21	-	<0.001	21	0	14	-	<0.001
		Cost	27	0	10	-5.004	<0.001	17	2	14	-	0.001
	P.M.	**Benefit**	30	1	6	-5.029	<0.001	11	2	23	-	0.022
		Risk	19	2	20	-	<0.001	24	2	11	-4.118	<0.001
		Cost	17	2	20	-	0.001	26	2	9	-4.341	<0.001
	2FA	**Benefit**	21	1	8	-	<0.001	7	6	15	-	1
		Risk	9	1	25	-	0.021	19	0	12	-	<0.001
		Cost	14	4	13	-	0.031	3	3	22	-	1
	Chg.P.	**Benefit**	29	1	2	-4.930	<0.001	13	12	9	-	1
		Risk	10	2	25	-	0.039	27	1	5	-4.725	<0.001
		Cost	25	1	10	-4.511	<0.001	16	0	16	-	<0.001
No Groups	Update	**Benefit**	12	1	10	-	0.003	12	4	8	-	0.077
		Risk	10	1	13	-	0.012	22	1	6	-	<0.001
		Cost	17	4	3	-	0.007	12	1	13	-	0.003
	P.M.	**Benefit**	19	4	9	-	0.003	20	3	8	-	<0.001
		Risk	21	0	10	-	<0.001	13	1	19	-	0.002
		Cost	21	5	7	-2.942	0.003	13	0	20	-	<0.001
	2FA	**Benefit**	15	2	9	-	0.002	14	6	5	-	0.115
		Risk	8	0	18	-	0.008	16	1	12	-	<0.001
		Cost	18	1	8	-	<0.001	4	3	19	-	1
	Chg.P.	**Benefit**	25	0	8	-	<0.001	23	3	6	-3.726	<0.001
		Risk	18	1	16	-	<0.001	32	0	6	-5.480	<0.001
		Cost	26	4	4	-3.834	<0.001	11	0	27	-	0.001

Table 9: Sign test results comparing *Individual* and *Social* ratings for each variable from tests performed on response sets separated by elements of advice and participants' group in the study (i.e., Yes or No group). For tests where there are fewer than 26 non-ties, the exact p is listed. In other cases, an asymptotic signifigance value is listed. Since Z statistics were not calculated for exact signifigance tests, this table only lists such a value where applicable.