You Want Me To Do What? A Design Study of Two-Factor Authentication Messages

Elissa M. Redmiles
University of Maryland
eredmiles@cs.umd.edu

Everest Liu
University of Maryland
eliu1234@umd.edu

Michelle L. Mazurek
University of Maryland
mmazurek@cs.umd.edu

ABSTRACT

Security messages that ask users to adopt new behaviors can be a crucial aspect of users' security decision-making. Prior work has focused extensively on how to design warning messages to discourage insecure practices. In this work, we instead examine how to design motivating security messages to encourage adoption, taking two-factor authentication (2FA) as a case study. To this end, we conduct an interview and participatory design study with 12 demographically diverse participants. Participants both critiqued existing 2FA messages and designed new ones. Drawing from the results of these interviews, we extract preliminary design options for authentication tool messages, which we plan to validate in future work.

1. INTRODUCTION

Over 80% of users report learning a security behavior from a security message [16]. These security messages typically ask users to adopt new behaviors, or warn them away from insecure practices. Given their prevalence, the effective design of these messages is crucial if we hope to keep users secure. Prior work has focused extensively on how best to design these messages to warn users away from insecure practices [4, 8, 9, 20]. Yet, significantly less work has focused on how best to design messages to motivate users to adopt a secure practice. These two goals, warning and motivating, are importantly different, as the mental processes that govern the avoidance of behavior differ significantly from those that govern the choice to adopt a new behavior.

In this work, we seek to better understand how to design motivating security messages for new authentication tools. Our work explores how best to design messages introducing two-factor authentication (2FA). 2FA is one of the security behaviors most recommended by security experts, but is insufficiently adopted by users: Ion et al. found in a large survey of experts that 74% recommend the use of 2FA, but only 40% of respondents in the Ion et al. survey reported using 2FA [13]. The messages used to invite users to enable 2FA may play a significant role in the adoption of this behavior, as our prior work finds that 41% of users report learning about 2FA based on a security message [16].

To this end, we conduct an interview and participatory design study with 12 demographically diverse participants. In each half-an-hour long interview, we asked participants to discuss existing 2FA invitation messages with us, and we also asked them to design a 2FA invitation message that they felt would motivate them to adopt the behavior. Drawing from the results of these interviews, we extract design options for authentication tool messages, compare our results to suggestions from prior work in usable security and the warning sciences, and present a prototype message based on these guidelines. Finally, we outline our plans for quantitative confirmatory work.

2. RELATED WORK

In this section we detail the prior work related to the design of security messages, including theoretical models for explaining user behavior and digital security-specific message design.

2.1 Models of User Behavior

Users often engage in practices that can seem illogical to an outsider. In the digital security domain, this is often attributed to a user's, perhaps misplaced, trust in the computer system to behave as expected in addition to a lack of understanding of the risks involved with their online behaviors [3]. Cranor et al. propose a model for user behavior in digital security to help explain these choices. The Human in The Loop model [6] describes the process for communicating digital security information to users as the following set of steps: (1) attention switch, (2) attention maintenance, (3), comprehension, (4) knowledge acquisition, (5) knowledge retention, and (6) knowledge transfer. In our work, we focus on steps 1 through 4, aiming to capture users attention and maintain it long enough for them to comprehend and acquire knowledge about 2FA in order to make an informed, and hopefully secure decision.

Relatedly, protection motivation theory (PMT), a general framework from the psychological sciences, explains users' behavior in response to threats. PMT suggests that behavior depends on four factors: perceived threat severity, perceived likelihood of threat occurring, perceived efficacy of preventative behavior, and perceive ability to protect themselves. Our work focuses on addressing the first three factors, with the intent to use design to improve perception accuracy.

2.2 Security-Specific Recommendations

Research from the warning sciences make a number of general suggestions for the design of product alert messages and warnings, in order to effectively capture people's attention or warn them away from danger. For example, showing the efficacy of making consequences explicit [15, 23], examining the cross-cultural impact of color on warning compliance [18], and emphasizing the utility of bullets for organizing complex information [17, 22]. Prior work in usable security has similarly studied a breadth of topics [4, 8, 9, 11, 20], including how best to communicate risk, the impact of message readability, the use of metaphors, and interactive and adaptive messages.

In the interest of brevity, we address just a subset of the findings here, focusing on recommendations for the visual design of messages as these are most closely related to our design- and text-centric work.

Egelman et al. examined phishing warnings; recommending on the basis of their findings that warnings should be distinguishable by severity, and also noting that it should be made difficult—for example by making users click through many screens—for users to proceed through a warning without first reading it [9]. Akhawe and Felt [2], on the other hand, focused their efforts on researching the effectiveness of browser security warnings. They analyzed many features that might effect warning success, including demographics, warning frequency, warning design, and warning complexity on the decisions of online users. Overall, they find that warnings have a significant effect on user security; contrasting with the prior belief that users did not read or heed such messages [3]. From their work, they distill a number of design guidelines including recommendations to avoid requiring users to click through many screens—a technique previously recommended to dissuade users from bypassing warnings. They also recommend not storing important information in "learn more" drop downs, as these items were rarely clicked.

Subsequently, Egelman and Schechter studied the effect of background color and text of a warning message on a user's decision to follow the warning [10]. They find that differences in background color and text varied the time users spent on a warning message, but did not significantly affect their decisions. Thus, they recommend focusing design attention on making risks salient through other mechanisms. Contrastingly, Weber et al. examined the use of physical security analogies and different colors to better emphasize the recommended options in each warning [21]. In our work, we draw inspiration from the participatory design workshops Weber et al. used to design the messages they evaluate.

Finally, Bravo-Lillo et al. explored the use of attractors to interact with the visual and textual information presented in a warning message [4]. They found that such attractors were effective and resilient to habituation, making them an attractive design element for inclusion in warning messages. It is unclear, however, what attractors should be used in the design of messages attempting to get users to engage in behaviors rather than avoid them. Should the user be forced to interact with key content, for example about the risks of choosing not to enable 2FA?

The aforementioned work focuses heavily on the design of security warning messages. On the contrary, our work is focused on encouraging users to engage in a new behavior, rather than warning them away. Thus, while we draw inspiration from the recommendations presented in this research, our work examines the importantly different question of how to design messages to motivate users to adopt new security behaviors for the prevention of potential future threats.

3. METHOD

We conducted an interview and design study with 12 participants in February 2017. Our protocol was approved by the University of Maryland IRB. In this section we describe our recruitment procedure, the details of our protocol, our analysis method and the limitations of our work.

3.1 Interview Protocol

In our study we sought to get feedback on four existing 2FA messages, in addition to soliciting new message designs from participants.

The four existing messages were from Google, Microsoft, Facebook, and Bank of America. These messages were selected as exemplars of a broad range of applications and message styles. In order to avoid participant fatigue, each participant critiqued only two 2FA messages. Each pair of messages was presented to two participants to minimize order effects. Appendix A shows each of the existing messages and their source; Table 1 shows which participants saw which messages and in what order.

Participants were first asked to describe 2FA as if they were explaining it to someone new to the technology. For those participants who did not recognize 2FA by name, we asked if they had ever needed to enter a code that was texted to their phone in order to login to a website; and then asked them to explain the purpose of this process. Those who did not recognize 2FA or could not accurately describe it were provided a simple definition.

Participants were then shown one message and then asked to describe how they felt about the message, what about the message would make them want to use or not use 2FA, how useful they felt that 2FA would be based on the message, concerns about 2FA, and what else they might like the message to tell them. Participants were then shown the next message and asked to compare this message to the prior message. They were then asked if they would be more or less likely to use 2FA if they saw this message or the last one. We followed up by asking the same questions asked about the first message they saw.

Finally, participants were engaged in a short participatory design task following a similar procedure to that used by Weber et al. [21]. Participants in our study were prompted as follows: "I would like you to design your own message to invite someone to use 2FA. While designing, please think about what would make you want to use 2FA? What information would you want to know? How would you want it to look?" Participants were provided with a large pad of paper and a 64-color pack of markers to construct their messages.

3.2 Recruitment

We recruited participants via Craigslist postings. We collected demographic information via a short screening survey and selected participants based on this information to ensure a maximally representative participant pool. Participants were compensated $15 for an approximately 30-minute study session.

3.3 Open-Coding Analysis

Two researchers used an open-coding process [19] to develop two codebooks: one based on three of the interviews and one based on three of the images. They then independently coded the remaining interview and image data with these respective codebooks. For the images, they achieved Krippendorff's alpha = 0.84, which is within the acceptable range for inter-rater reliability. For the interview transcripts, they achieved alpha = 0.79, which is acceptable for an exploratory study such as this one.

ID	Sex	Age	Race	Educ.	Income	Msg.
P1	F	18-29	A	H.S.	<$30k	M1, M2
P2	M	18-29	W	S.C.	$75-$99k	M3, M1
P3	F	50-59	B	H.S.	<$30k	M2, M1
P4	M	18-29	W	S.C.	<$30k	M4, M2
P5	M	Over 70	W	B.S.	$100-$125k	M2, M4
P6	F	30-39	W	A.D.	$75-$99k	M2, M3
P7	M	30-39	H	B.S.	$30-$50k	M1, M4
P8	M	40-49	B	B.S.	$50-$75k	M4, M3
P9	F	40-49	B	A.D.	<$30k	M3, M2
P10	F	30-39	A	A.D.	$75-$99k	M3, M4
P11	M	50-59	B	B.S.	$75-$99k	M1, M3
P12	F	40-49	W	H.S.	<$30k	M4, M1

Table 1: Participant's demographics and the messages they were shown, in order. Education abbreviations: high school graduate (H.S.), some college (S.C.), Bachelors degree (B.S.), and advanced degree (A.D.).

3.4 Limitations

As with all qualitative studies, our work is limited by the size and diversity of our sample. We followed recommendations from prior work to interview 12-20 participants until new themes stopped emerging [5], and we attempted to recruit as diverse a sample as possible in order to maximize the generalizability of our results. Additionally, our work may be limited by interviewer-induced bias: participants may have felt a desire to respond in certain ways in order to garner the interviewer's approval, and/or the interviewer's age, gender, or race may have influenced participants' responses. Because we did not ask about participants' own security behaviors, and made clear to participants that they were not evaluating the researchers' own work, we hope to have mitigated some of these biases. Further, we hope that by recruiting a diverse sample we increased the likelihood that relevant themes would be mentioned by at least one participant.

4. RESULTS

In this section we present the results of our analysis of participants' critiques of existing 2FA messages and participants' designs of new 2FA messages.

4.1 Participants

The demographics of our participants nearly match those of our demographic area with regard to race and gender [1]. Our participants are slightly more educated than the D.C metro area with 58% of our participants holding at least a B.S. compared to 44% of D.C.-metro area residents having the same educational attainment. That said, our participants are also less wealthy than the D.C. metropolitan area, with 91% of our participants making under $100K as compared to 60% of the D.C. area. Our participants are also differed in their age distribution, with 50% of our participants being between the ages of 29 and 49 compared to 28% of the D.C.-metro population. Table 1 shows a each participant's demographics.

To evaluate participants' baseline understanding and experience with 2FA, we asked them to describe it (see Section 3.4 for more details). Most (9) participants described 2FA as a security measure or a way of verifying your identity. For the three participants who did not recognize 2FA by name or description, or did not accurately describe it, we provided a simple explanation.

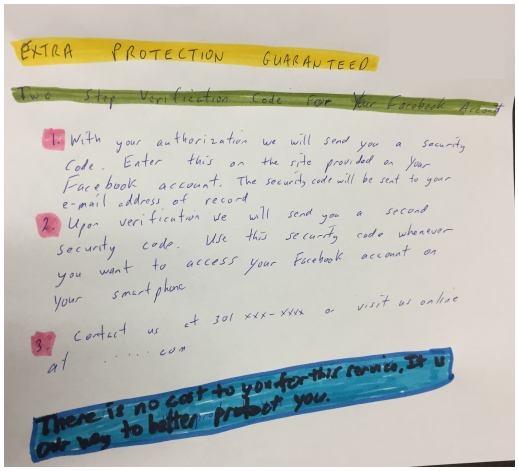

Figure 1: P5 draws a message with bullets outlining how 2FA will work and the steps to set it up.

4.2 User Preferences for 2FA Messages

Overall, five participants preferred M2, five preferred M1, one preferred M3, and one had no preference (see messages in Appendix A). Below we present the results of our open-coding analysis, first detailing participants' preferences for 2FA message content and the describing their UI preferences.

4.3 Feedback on Message Content

Complexity and Clarity. Eight participants reported finding the messages confusing. P5 said, "I thought it was a little confusing since I'm not technologically skilled," and similarly P4 said, "I don't know what they're really asking for here." Similarly, six participants mentioned liking a message because it was simple (these six participants preferred either M1 and M2 over M3 or M4). P4 says about M2, "this one just explains the two step verification. It's more clear about why they're doing it and makes you feel more protected because it helps explain that they want two different forms of verification." Seven participants also mentioned disliking messages that were vague, and four mentioned preferring messages that were detailed. They were especially interested in messages that specifically described either what 2FA was or its purpose. They especially preferred simple, direct explanations: P8 says "I like [this message] because it's fairly simple in it's explanations. [In this other one] there is no real explanation... it just says 'a second layer of protection'." Similarly, P7 compares M1 to M4, "[M1] kind of explains a little bit more what [2FA] does, while [M4] just tells you what you can do with it. I don't know, I feel like I have a little bit more information here [in M1], like, why is it that I'm doing this."

Mirroring this critique feedback, when designing their messages, nine participants included a clear explanation on how to enable 2FA, in contrast to only one of the existing messages including such clear instructions. Eight of these participants used bullet points or steps to reduce confusion in their messages, while bullet points were only included in one existing message (M2). See Figure 4.3 for an example of how P5 explained the steps required to set up and use 2FA. P12 also included step-by-step instructions, explaining that they did this "so [that] when [the user] decides [they] want to keep [their] account secure then you just tell them how easy it is to do it."

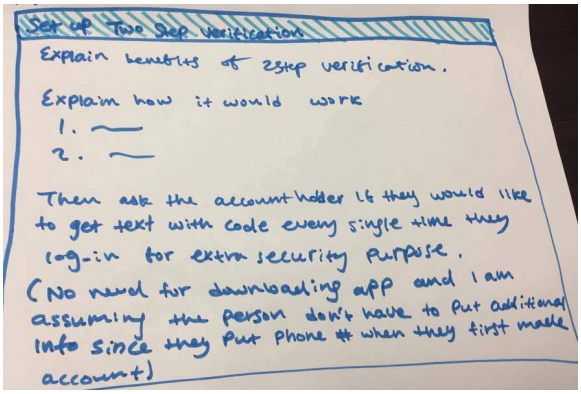

Figure 2: P1 draws an all blue message explaining why and how to use 2FA.

Secure or Protect? In their critiques, eight participants mention preferring messages that mention security (or disliking that a message did not mention security); six of these participants plus two additional participants also used "security" or "secure" when sketching their 2FA designs. This is in contrast to five participants who used the word "protect" (one participant used both protect and secure). On the other hand, two of the existing messages include only the word protect, one includes both secure and protect, and one includes only secure. P9 explains their preference for the word secure as follows: "The fact that it says, adds an extra layer of security. [In this other message] there's no mention of security: it says, 'we're here to help you protect your account', but, I would assume that they're already protecting my account. I guess the part that says extra layer of security is what would convince me."

Information about Cost and Time. Many of participants' critiques focused not on the messages, but on 2FA itself. Participants were concerned about whether 2FA was free (2 of 12) and about how long it would take them to set up and log in (6 of 12). They advocated for including information about costs or time-lengths in the message text. Similarly, three participants mention that whether they will enable 2FA is tied less to the message and more to the type of account. For example, P6 says, "maybe [I would use it], for bank accounts, or something like that it'd be worthwhile, but most of the things I do online, Facebook, email, there's nothing interesting in those accounts so there's no reason to have an extra layer of protection on those." These three participants all note that accounts other than their bank account are "not important," suggesting that users may underestimate the risk a compromised email account may pose [12]. We hypothesize that including information in 2FA invitation messages about how accounts other than bank accounts can have financial implications may increase adoption.

Privacy. Finally, six participants mentioned that the messages should address privacy concerns and how personal information provided to set-up two factor authentication will be used. These concerns were tied to account types for four of the participants: P7 says "I feel like Facebook has started to become very, very invasive.... It's been changing rules and the way they do things. Basically, I'm afraid now of giving them any type of information that at this stage might be used for this two step verification process, but maybe later on will be used for something else.... I might be a little bit more lenient with Gmail, but still." P10 suggests that messages should include privacy-related information in a learn more section.

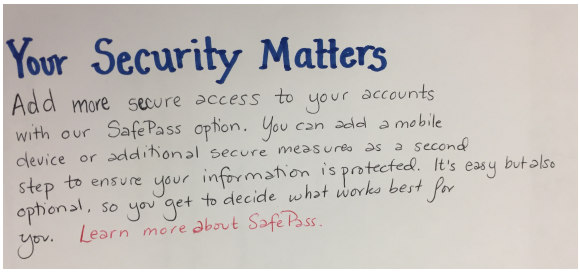

Figure 3: P10 draws a message with the personalized headline "Your Security Matters."

4.4 Feedback on Message Appearance

Blue. We find that the majority of participants (8 of 12) chose to use the color blue when sketching their messages, and five of those who did explicitly noted that they used blue because it felt "less scary" than other colors. Figure 4.3 shows an example all-blue message; when describing this message that they drew, P1 said "I would use blue, because red just intimidates me and red is more like 'Alert, alert. Something's going wrong'." Similarly, P6 said they used blue instead of red, "because red's a little scary."

Graphics. Four participants mentioned disliking the graphics included in the existing messages. They noted that the graphics made them feel that the message was "not serious" (P6). P5 says "The phone icon [in M4] turned me off, really, right away... it gives me a concern that this may not be legitimate," while critiquing M4. Indeed, in their designs, participants avoided the use of graphics, with only three of 12 participants creating designs that included graphics. P10, who chose not to use a graphic in their message (Figure 4.3 said "sometimes I feel like it feels a little bit more formal when there's less pictures ... it feels more serious." This infrequent use of graphics is in contrast to the existing messages, three of which include graphics. Participants also avoided creating interactive messages (two participants), potentially because interaction is hard to imagine in paper sketches, or because such interaction would consume more time and thus is undesirable as suggested by related work [2]. Finally, only four participants included reasons why the user should enable two factor authentication.We hypothesize that this may relate to users' strong concerns about how to enable 2FA, which took precedence over why they should do so; a

Addressing the User Directly. Lastly, second-person, personal headlines were used heavily in participants' sketches, with eight participants including headlines such as "Your Security Matters!" as shown in Figure 4.3. When describing their message, P10 says they used this personal headline because that way, "this message tells me, hey, we care about you. That [makes it] important to me." They compared such personal headlines to those in the existing 2FA messages, which were more general and passive, and did not refer to the company helping the user or the user helping themselves.

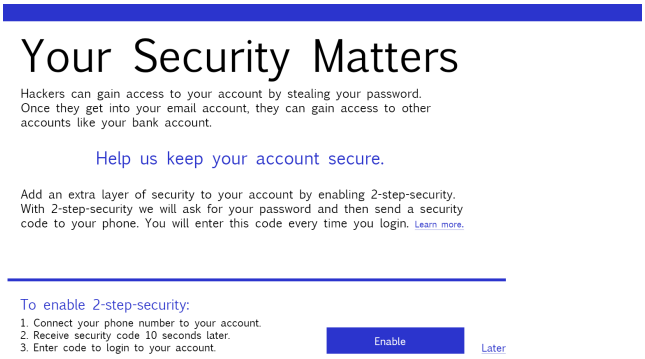

Figure 4: Prototype message design.

5. DISCUSSION AND FUTURE WORK

In this section we distill a set of recommendations for the authentication tool invitation messages and outline our plans for confirmatory work.

5.1 Design Options

Below, we present preliminary design options that can be explored in future work for the development of security messages aimed at introducing new authentication behaviors to users and we discuss these recommendations in the context of prior work.

- **Use Blue.** Our participants strongly favored the use of the color blue, which prior work has shown is perceived as peaceful and calm [18], in the design of their messages, and avoided red, which our participants describe as signaling something scary. User preference for blue in this study differs from prior work on SSL warnings, which found that PD participants favored signal colors in warning design such as red and green [21]. We hypothesize this may be due to the difference in goals between 2FA messages and SSL warnings—the former seeks to have participants click through to adopt a behavior, while the later aims to have them click away to avoid a dangerous behavior.

- **Use Personal Headlines.** Eight of our twelve participants used personal headlines in their designs of 2FA messages. These headlines made participants feel that the company presenting the message cared about them and that using 2FA must thus be in their best interest. This is a finding not presented in prior work, and which we hypothesize may be useful in communicating bi-directional investment, helping the user not feel over-burdened by a security task they may not otherwise wish to perform.

- **Use Bullet Points** Eight of our participants used bullet points in their designs, typically to provide a step-by-step explanation of how to set up two factor authentication. Participants noted in their critiques of existing message designs that they were hesitant to set up new authentication tools such as 2FA if they did not understand up front what would be required of them. Their desire to use bullet points in these explanations aligns with recommendations from the warning sciences [22], which suggest the use of bullet points to explain unfamiliar or complex steps [17].

- **Avoid Graphics** Surprisingly, our participants overwhelmingly (9 of 12) chose not to include graphic elements in their 2FA invitation messages. They cited feeling that these graphics made the message less serious and more suspicious. This finding is new to usable security but mirrors findings from the warning sciences that icons may not significantly enhance messages [7,14]. We caveat this finding by noting that our sample did not include any participants with low English literacy; graphics may be more useful for other populations not examined in this work.

In Figure 5 we present a prototype message design built upon the existing messages that participants most preferred (M1 and M2) and based on the design guidelines we distilled from our findings; this design was reviewed by two HCI experts to ensure following of design best practices.

5.2 Planned Future Work

One guideline we were surprised was not suggested by our results was the inclusion of information about consequences or risk in the messages. Only 4 participants placed a "why" statement related to the necessity of 2FA or consequences that might be incurred from not using it. This is in contrast to prior work [9, 15] which found that the inclusion of such consequences can be a powerful behavior motivator. It is possible that participants did not include these why elements in their designs because they were not comfortable enough with the content area to feel comfortable describing the risks. Thus, our confirmatory work will also explore the inclusion of specific risks and consequences in authentication invitation messages.

6. ACKNOWLEDGEMENTS

We thank Rachael Marr, Jennifer Cowley, and Greg Shannon for their help developing this study. This work was supported in part by Maryland Procurement Office contract no. H98230-14-C-013.

7. REFERENCES

[1] U.s. census bureau (2015). american community survey 1-year estimates. washington- baltimore-arlington.

[2] D. Akhawe and A. P. Felt. Alice in warningland: A large-scale field study of browser security warning effectiveness. In *USENIX Sec.*, 2013.

[3] C. Bravo-Lillo, L. F. Cranor, J. Downs, and S. Komanduri. Bridging the gap in computer security warnings: A mental model approach. *IEEE S&P*, 2011.

[4] C. Bravo-Lillo, S. Komanduri, L. F. Cranor, R. W. Reeder, M. Sleeper, J. Downs, and S. Schechter. Your attention please: Designing security-decision UIs to make genuine risks harder to ignore. In *SOUPS*, 2013.

[5] K. Charmaz. *Constructing grounded theory : a practical guide through qualitative analysis*. Sage, 2006.

[6] L. F. Cranor. A framework for reasoning about the human in the loop. *UPSEC*, 2008.

[7] S. Davies, H. Haines, B. Norris, and J. R. Wilson. Safety pictograms: are they getting the message across? *Applied ergonomics*, 1998.

[8] F. De Keukelaere, S. Yoshihama, S. Trent, Y. Zhang, L. Luo, and M. Zurko. Adaptive security dialogs for improved security behavior of users. *Human-Computer Interaction–INTERACT 2009*, pages 510–523, 2009.

[9] S. Egelman, L. F. Cranor, and J. Hong. You've been warned: An empirical study of the effectiveness of web browser phishing warnings. In *CHI*, 2008.
[10] S. Egelman and S. Schechter. The importance of being earnest [in security warnings]. In *FC*, 2013.
[11] M. Harbach, S. Fahl, P. Yakovleva, and M. Smith. Sorry, i don't get it: An analysis of warning message texts. In *FC*, 2013.
[12] M. Honan. How apple and amazon security flaws led to my epic hacking, 2012.
[13] I. Ion, R. Reeder, and S. Consolvo. "...no one can hack my mind": Comparing expert and non-expert security practices. In *SOUPS*, 2015.
[14] L. S. Jaynes and D. B. Boles. The effect of symbols on warning compliance. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 1990.
[15] K. R. Laughery, K. P. Vaubel, S. L. Young, J. W. Brelsford, and A. L. Rowe. Explicitness of consequence information in warnings. *Safety science*, 16, 1993.
[16] E. Redmiles, S. Kross, and M. L. Mazurek. How i learned to be secure: a census-representative survey of security advice sources and behavior. In *CCS*, 2016.
[17] A. Sears and J. A. Jacko. *Human-computer interaction: design issues, solutions, and applications*. CRC Press, 2009.
[18] M. Silic, D. Silic, and G. Oblakovic. The effects of colour on users? Compliance with warning banner messages across cultures. ECIS, 2016.
[19] A. Strauss and J. Corbin. *Basics of qualitative research: Procedures and techniques for developing grounded theory*. 1998.
[20] J. Sunshine, S. Egelman, H. Almuhimedi, N. Atri, and L. F. Cranor. Crying wolf: An empirical study of ssl warning effectiveness. In *USENIX Sec.*, 2009.
[21] S. Weber, M. Harbach, and M. Smith. Participatory design for security-related user interfaces. *Proc. USEC*, 15, 2015.
[22] M. Wogalter. Purposes and scope of warnings. *Handbook of Warnings (3–9); Wogalter, M., Ed*, 2006.
[23] M. S. Wogalter, V. C. Conzola, and T. L. Smith-Jackson. Research-based guidelines for warning design and evaluation. *Applied ergonomics*, 2002.

APPENDIX

A. EXISTING MESSAGES CRITIQUED BY PARTICIPANTS

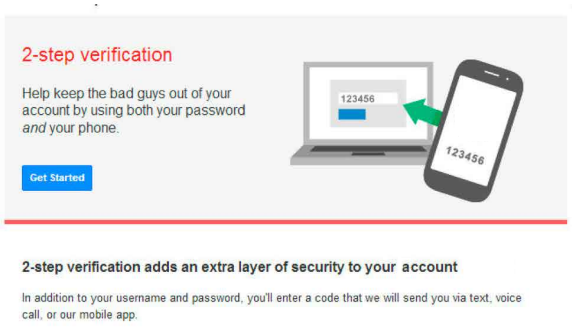

Figure 5: Google's 2FA invitation message (M1)

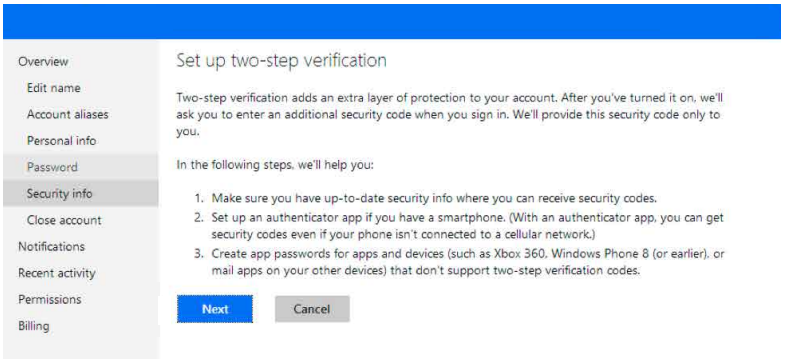

Figure 6: Microsoft's 2FA invitation message (M2)

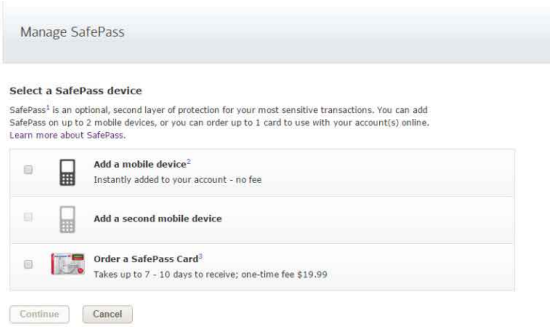

Figure 7: Bank of America's 2FA message (M3)

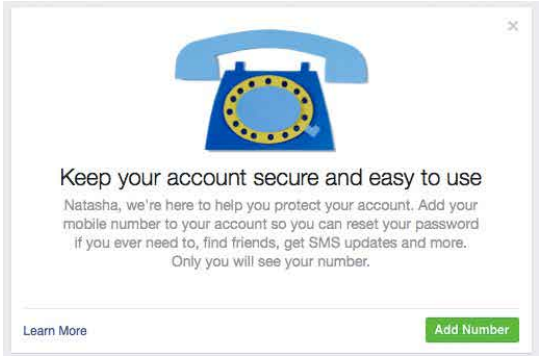

Figure 8: Facebook's 2FA invitation message (M4)