Divergences in Blame Attribution after a Security Breach based on Compliance Behavior: Implications for Post-breach Risk Communication

Ehsan Ul Haque
ehsan.ul_haque@uconn.edu
University of Connecticut
Storrs, Connecticut, United States

Mohammad Maifi Hasan Khan
mohammad.khan@uconn.edu
University of Connecticut
Storrs, Connecticut, United States

Md Abdullah Al Fahim
md.fahim@uconn.edu
University of Connecticut
Storrs, Connecticut, United States

Theodore Jensen
theodore.jensen@uconn.edu
University of Connecticut
Storrs, Connecticut, United States

ABSTRACT

"Attribution of self-blame" is a spontaneous affective and cognitive self-evaluative reaction and is an important predictor of proactive and positive coping response behavior after a negative event. While blame attribution can indeed affect the efficacy of post-breach communication and subsequent behavior in the cybersecurity context, it is not clear whether and how the process of blame attribution may vary in post-data breach contexts based on users' past security compliance behavior, which can inform the design of post-breach risk communication. As such, to examine the process of blame attribution after a data breach, we run a 2 (user type) x 2 (account type) x 2 (usage scenario) between-group study on Amazon's MTurk platform. The vignette-based study scenario incorporates multiple stakeholders who may share the responsibility for the data breach (e.g., the negligent user (Bob), the software company holding the data, and the attacker). From the analysis of 255 participant data, we discover that adopters are more likely to hold the protagonist of the story responsible due to his negligent behavior. In contrast, non-adopters are more likely to hold the external entity, such as the service provider, accountable for the account compromise, exhibiting "defensive attribution" of blame. Results further indicate that account types and usage scenarios affect how blame is distributed among different entities. The implications of our findings for effective pre and post-breach risk communication are discussed in the paper.

CCS CONCEPTS

• **Human-centered computing** → **Empirical studies in HCI**; • **Security and privacy** → **Usability in security and privacy**.

KEYWORDS

Data breach, Usable cybersecurity, Risk communication, Blame attribution

ACM Reference Format:
Ehsan Ul Haque, Mohammad Maifi Hasan Khan, Md Abdullah Al Fahim, and Theodore Jensen. 2023. Divergences in Blame Attribution after a Security Breach based on Compliance Behavior: Implications for Post-breach Risk Communication. In *The 2023 European Symposium on Usable Security (EuroUSEC 2023), October 16–17, 2023, Copenhagen, Denmark*. ACM, New York, NY, USA, 21 pages. https://doi.org/10.1145/3617072.3617117

EuroUSEC 2023, October 16–17, 2023, Copenhagen, Denmark

ACM ISBN 979-8-4007-0814-5/23/10…$15.00
https://doi.org/10.1145/3617072.3617117

1 INTRODUCTION

Equifax data breach is estimated to have compromised over 146.6 million consumer accounts [82], and exposed sensitive information such as names, date of birth, social security numbers, credit card information, and driver's license numbers of approximately 209,000 consumers [34]. Surprisingly, Zou et al. found that, even though participants knew and showed concerns about the breach, many did not take any protective measures afterwards [82].

While a significant volume of prior research noted that gaps in risk perceptions[14, 67], lack of self-efficacy [74, 81], and inconsistent/erroneous mental models [3, 73, 76] may explain users' noncompliant behavior, only a few looked into factors that may explain users' post-breach lack of action. Notably, the handful of works that looked into users' post-security breach behavior primarily focused on the negative effects of security breach on users' trust and possible trust repair strategies rather than how the breach may affect users' future security behavior [7, 9, 52, 68]. While trust repair is essential following a security breach, it mainly focuses on service providers' interests and well-being (i.e., mechanisms to regain trust and maintain the user base) and does not shed light on users' post-breach coping response behavior.

Although not widely explored in cybersecurity, prior efforts in different domains looked into users' coping response behavior following negative incidents with varying degrees of harm (i.e., monetary loss to serious bodily injuries) and identified blame attribution as a predictor of coping response behavior [17, 55, 56, 65]. As per the attribution theory, blame attribution refers to the human tendency to look for plausible explanations, especially after a negative event [39, 77], and depends on the perception of causality of the event. In this vein, prior efforts in psychology and healthcare

have investigated and acknowledged the influence of factors such as self-blame and self-responsibility on users' decision-making process [8, 17, 44, 45, 56, 66]. In addition, literature focusing on social cognitive frameworks has confirmed the relationship between attribution of self-blame and positive coping response behavior [17, 66] and proactive behavior [44].

Given the recognition of blame attribution in behavior adoption and positive coping response, we argue that it is crucial to study and understand the process of blame attribution following security breaches, which can inform the design of post-breach communication promoting future security measures (e.g., change of passwords, enabling of two-factor authentication). Further, it is important to investigate whether past security compliance behavior (i.e., user type: adopters vs. non-adopters) affects post-breach blame attribution differently or not. If such differences exist among adopting and non-adopting populations, post-breach communication must account for that and tailor messaging accordingly. Motivated by prior efforts on blame attribution and its influence on coping response behavior, in this paper, we investigate how adopters and non-adopters of a security measure (e.g., 2-FA) differ in attributing blame to the entities who may share the responsibility of a data breach.

In addition to past security compliance behavior, we also consider account contexts, as prior efforts noted that the perceived value of an object/account can influence users' protective behavior. For instance, prior work in screen locking showed that a higher perceived valuation of data stored on the phone can motivate users to adopt security features [5]. A similar result is reported in the context of 2-FA, where high-income users were found to be more likely to adopt 2-FA due to higher perceived financial risks due to cyberattacks [58]. This is supported by loss aversion theory, which suggests that people put more effort into avoiding losses than an equivalent gain [48]. Based on these prior insights, we hypothesize that contexts such as account type (i.e., personal vs. official), which can affect the perceived value of the data stored in the account, may play an essential role in blame attribution and incorporate in the study as a variable.

Finally, we expected that an email account that is being used for both business and personal purposes (i.e., account usage: combined) in comparison to one that is being used for only intended purposes (i.e., account usage: intended) could affect the process of blame attribution. Specifically, using a personal account for both official and personal purposes (or an official account for both official and personal purposes) may be perceived as negligent behavior, which can affect the blame attribution. Hence, we also included the account usage factor in the study. These considerations lead us to look at blame attribution across user groups and account/usage contexts by asking the following research questions:

RQ1 (a) How do adopters and non-adopters of a security tool/feature differ in attributing blame after a data breach? (b) What reasonings are given to justify non-adopting behaviors that may be used to deflect blame?

RQ2 How does account type (e.g., compromised account in context) affect the attribution of blame after a data breach? How does user type (adopters vs. no-adopters) impact the effect of account type on blame attribution?

RQ3 How does account usage scenario affect the delegation of blame after an attack? Is there any interaction effect of user type, account type, and account usage scenario on blame attribution?

To investigate these research questions, we conducted a 2 (user type: adopters vs. non-adopters) x 2 (account type: official vs. personal email) x 2 (usage scenario: intended vs. combined (both official and personal)) between group study. We collected data from 255 participants across eight groups who were asked to attribute blame towards a hypothetical character, Bob, who, despite being prompted, did not adopt 2-FA for his email account and eventually got compromised by a data breach. Our investigation makes the following key contributions:

- We uncover key differences between the adopting and non-adopting population of 2-FA regarding the attribution of blame in a data breach scenario and how blame attribution is affected by account types and usage scenarios.
- Leveraging quantitative and qualitative data analysis, we document possible reasonings behind the differences in blame attribution across groups, which shed light on the defensive attribution of blame.
- We discuss our results' implications for pre and post-breach cybersecurity risk communication.

2 RELATED WORK

2.1 Divergences in Security Behavior

A large volume of prior effort exists that looked at different theoretical frameworks such as Protection Motivation Theory (PMT) [62], Theory of Planned Behavior (TPB) [2], and Health Belief Model (HBM) [21] to explain and alter users' cybersecurity behavior in various contexts [4, 5, 20, 37, 50, 80]. Towards understanding users' cybersecurity behavior, researchers have identified substantial differences in mental models and adoption behavior among different populations (e.g., expert vs. non-expert, adopters vs. non-adopters) [19, 42, 49, 72]. A number of efforts attempted to identify the factors that may explain the divergences in security behavior among different populations. Among multiple factors, risk perception - a subjective judgment made by users based on the severity of risks - is identified as one of the critical factors in the literature that can explain (to some extent) and lead towards changes in behavior [6, 29, 35, 37, 41, 49, 59, 67]. Along this line, various research investigated the effect of risk communication messages in altering risk perceptions and ultimately adopting the target behavior. For example, Albayram et al. designed a video incorporating fear appeal in intervention messages and demonstrated the effectiveness in screen locking behavior [5]. Harbach et al. and Das et al. also reported similar results using the concept of risk alteration [23, 36], underscoring the importance of considering risk perception in cybersecurity behavior.

While these prior efforts identified factors that may explain the divergence in behaviors, another group of work noted that adequate understanding of risks and safe behavior does not always lead to compliance [40, 83]. Further, no direct correlation is found between participants' technical background and their actions to control their privacy [49], underscoring the nuanced context-sensitive nature of security behavior and decision-making.

2.2 Personal Responsibility and Attribution of Blame

Attribution of blame refers to the human tendency to look for plausible explanations, especially after a negative event [39, 77]. According to attribution theory, reaction (e.g., blame) after an (adverse) event depends on the event's perceived causality. It can be influenced by four causal dimensions, namely, *locus of causality* (i.e., perception of whether the cause of the event is within the scope of the entity or outside the scope of the entity), *stability* (i.e., perception of whether the cause of the event is stable or may change with time), *controllability* (i.e., perception of whether the cause of the event was within the control of the entity or not), and *intentionality* (i.e., perception of whether the cause of the event was intentional by the entity or not) [38].

An interesting observation related to attribution of blame and responsibility was pointed out by Shaver, who noted that attribution of responsibility and blame depends on observers' perception of similarity with the offender responsible for the adverse incident, which he termed as *Defensive Attribution* [65]. According to the defensive attribution theory, when the similarity between the observer and the offender is high, the observer tends to attribute lower blame to the offender and show higher leniency in judgment. When the similarity is perceived as low, the observer will likely blame the offender more. When an observer perceives themselves to have a similar personality to an offender, in that case, the defensive attribution is evoked as the perceived similarity emphasizes a chance of putting themselves in a similar situation as the offender [64, 65]. In such a situation, the observer tends to feel that the circumstance was *unfortunate*, attributing lower blame to the offender. Shaver further noted that personal similarities like age could also evoke defensive attribution. Other factors, such as ethnic and sociocultural similarities between the observer and offender, were also shown to evoke defensive attribution [26, 64, 69].

The attribution of blame and its effect on subsequent behavior is also studied in the context of organizational settings. In organizational contexts, blame attribution is viewed from the perspective of employees' psychological contract breach (PCB) [1, 22]. Here, PCB refers to the cognitive perception of digression from the organization's promises (formally or informally) to its employees [55]. As a result of such digression, employees feel betrayed, which further affects their attitude and performance towards the organization [55].

We argue that blame attribution is relevant to cybersecurity behavior where users are likely to have certain expectations (e.g., quality of service, security) from service providers and may feel betrayed when the service gets compromised. Such adverse incidents will likely cause users to delegate blame among the responsible parties following a negative event [79]. To shed light on how end-users feel about who should be accountable for user security, a survey of over 9000 participants indicated that consumers mostly disagree that end-users should be responsible for protecting their data [71]. Gemalto's report echoed these findings where 70% participants felt that the responsibility of securing end-users lies with the organizations and service providers [32]. Peck et al.'s work further supported the results where participants mostly held the service providers responsible for a hypothetical data breach scenario [57].

While prior efforts pointed out users' tendency to blame the service providers, they did not examine whether such delegation of blame is affected based on users' past compliance behavior. Further, as several factors and entities can contribute to an unintended occurrence, blaming a single party in a complex, interrelated system can be challenging and requires careful investigation to understand the process of blame attribution [24, 25].

3 METHODOLOGY

Our study had three independent variables, each with two levels. Our first independent variable, "user type," includes adopters and non-adopters of 2-factor authentication (2-FA). In the study, we defined adopters as those using 2-FA for at least one of their email accounts. As part of our prescreening questionnaire, we asked participants - "Do you currently use two-factor authentication (a.k.a. 2-FA or two-step verification) for any of your email accounts (e.g., official, personal, or business email accounts, etc.)?" Participants who responded "yes" to this question were assigned to one of the adopter groups; otherwise, they were assigned to the non-adopter groups.

The second independent variable, "account type" (i.e., official vs. personal email account), indicates the type of email account compromised in the hypothetical attack scenarios presented in the vignettes.

Our third independent variable "usage scenario" includes two levels: intended use - when Bob (i.e., the hypothetical character) used his official email for official purposes or personal email for personal purposes only, vs. combined use - when Bob used his email for both official and personal purposes irrespective of the account type. These led to 8 groups for the between-subject study (Table 1).

Our dependent variable for the study is the attribution of blame among the stakeholders of the hypothetical breach incident presented in the vignettes. The blame attribution items are 7-point Likert scale questions (Strongly disagree to Strongly agree) in the format "Bob should blame *stakeholder* for *reason*". The stakeholders are the parties who could have played a role in preventing the breach incident (i.e., Bob himself, the company/provider in the context, the attacker, and the passive influencing entities such as Bob's friends/families, the government, and the 2-FA promotional message).

3.1 Design of the Vignettes

For different levels of our independent variables, we had a set of four vignettes that presented Bob's scenarios before the attack (part 1) and another two vignettes for scenarios after the attack (part 2).

In part 1 of the vignettes, participants were introduced to Bob. They were informed that he worked for a health insurance company and had an official email account provided by his company. For groups 1 and 2, Bob used his official email account for official purposes only (intended use - Group 1A, 1B) or his official email account for both official and personal purposes (combined use - Group 2A, 2B). For groups 3 and 4, alongside official email, Bob also had a personal email account of his own, which he used either for personal purposes only (intended use - Group 3A, 3B) or both official and personal purposes (combined use - Group 4A, 4B).

Account Type	Usage Scenario	User Type	Groups
Official Email Account	Intended use (official purposes)	Adopter	1A
		Non-Adopter	1B
	Both official and personal purposes	Adopter	2A
		Non-Adopter	2B
Personal Email Account	Intended use (personal purposes)	Adopter	3A
		Non-Adopter	3B
	Both official and personal purposes	Adopter	4A
		Non-Adopter	4B

Table 1: Study groups.

The storylines for part 1 vignettes further showed that Bob received a message promoting a feature called 2-FA from either his company for his official email account (Group 1A, 1B, 2A, 2B) or the personal email service provider for his personal email account (Group 3A, 3B, 4A, 4B). In our case, the company or the personal email service provider was the entity in the context that provided the email service to Bob.

We did not use any specific name (e.g., Gmail) for the email service provider to avoid personal bias towards any specific service provider (either favorable or unfavorable). Part 1 of the storylines ended by informing participants that Bob did not enable the 2-FA feature upon receiving the 2-FA promotional email. Storylines are attached in Appendix A.2.

We used Google's official 2-FA promotional message during the study and edited out any wordings or distinctive graphics that may direct participants to identify the source. This design choice was made based on prior work that noted that brand trust crucially impacts users' security and privacy-related behavior [53]. Using the anonymized version of Google's official 2-FA image allowed us to show participants an industry-standard image containing the elements that may promote behavior change (e.g., risk, self-efficacy, fear appeal) while at the same time making it harder for the participants to guess the true origin of the image to keep it consistent with the concept of an unknown, generic entity. Original and edited versions of the 2-FA promotion messages are attached in Appendix A.6. Note that the current version of Google's 2-FA image has been updated since this study was performed. However, as the goal of our study is not to test the effect of different message framing on blame attribution, such changes made by Google (which is expected and beyond our control) do not affect the findings of our study.

To ensure that the manipulation was successful and participants did not recognize the origin of the image, we asked participants two questions after they saw the vignette part 1 - (a) "Have you ever seen the particular image displayed in the storyline before?", and (b) "Where did you see this image?". We removed participants who answered "yes" to the first question, irrespective of their qualitative response to the second question.

After participants viewed the image, part 2 of the vignettes informed participants that attackers broke into either the company email authentication server (groups 1 and 2) or the personal email service provider's authentication server (groups 3 and 4) and stole login credentials of several thousand employees (groups 1 and 2)/users (groups 3 and 4) including Bob's. Storylines for part 2 further added that Bob was unaware of the attack. The following day, attackers used Bob's credentials to log into Bob's official email account (groups 1 and 2) or personal email account (groups 3 and 4). Part 2 of the storylines ended with informing participants that attackers changed Bob's email passwords, preventing him from logging in to his account.

Two parts of the vignette were used to observe participants' attitudes towards Bob under both pre and post-breach scenarios. Specifically, after part 1 of the vignette (where participants were informed that Bob received the 2-FA notification email and decided not to enable it), we asked participants whether they thought Bob did the right thing by not enabling 2-FA and why. We asked this before participants watched part 2, where they were informed of the attack. This allowed us to investigate participants' reasonings about both for and against Bob's decision of not enabling 2-FA without being biased by the consequence (i.e., security breach) of Bob's negligent action. Questions after Part 2 allowed us to investigate their process of blame attribution. Without having a 2 stage design, we would not be able to investigate participants' reasonings for/against insecure behavior without possible confounding effects due to the security breach depicted in part 2.

3.2 Participant Recruitment

We recruited participants from Amazon's Mechanical Turk (MTurk) platform. We restricted participation in the study to those at least 18 years of age and currently living in the United States. Furthermore, to ensure data quality, we recruited participants who had completed at least 1000 Human Intelligence Tasks (HIT) and had a HIT approval rate of 95% based on recommendations from prior works [70]. We restricted participants to those who use a Windows personal computer to avoid possible confounding effects due to different operating systems [13]. Further, we restricted to those who use Gmail as one of the email account(s) for critical tasks (e.g., official business, online bill payment, etc.).

3.3 Attention and Manipulation Checks

To ensure data quality, we added four attention check questions (in the form, 'Please select "somewhat disagree" for this statement.') at random places throughout the survey. We only considered responses as valid that passed all four of these attention check questions. In addition, to ensure that participants read and understood the storylines correctly, we asked multiple manipulation check questions about the storyline after they were presented with the

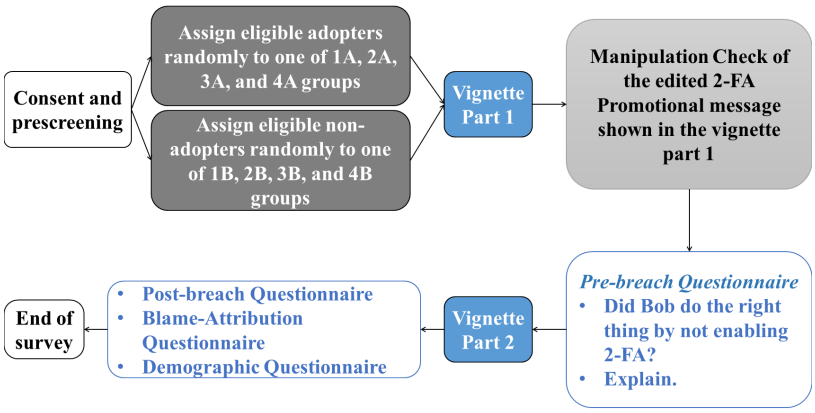


Figure 1: Study Procedure.

vignette. Participants needed to answer these questions correctly before proceeding to the next part of the survey. They had the opportunity to go back and reread the storylines to be able to answer all of the questions correctly.

3.4 Survey Flow

After giving consent, participants answered a prescreening questionnaire. Based on their responses, eligible participants were assigned to one of the eight groups and were redirected to part 1 of the storyline for their assigned group. Participants then answered the questionnaire about the storyline, the 2-FA image, and other Likert scale items based on the storyline. Then, participants were shown part 2 of the vignette. After answering the second set of attention check questions related to the part 2 storyline, participants answered items related to blame attribution and the demographic questionnaire. The survey took about 30 minutes to complete ($M = 32.87, SD = 18.58$), and eligible participants were compensated $3 for completing the survey. The survey was hosted on the university's Qualtrics server and was approved by the university Institutional Review Board (IRB). The survey flow is presented in Figure 1.

3.5 Survey Data Analysis

We recorded a total of 399 responses for the main survey. Out of the 399 responses, 114 (28.57%) participants said they had seen the 2-FA promotional image before. In addition, 29 (7.3%) participants failed at least one of our survey's four attention check questions. We removed these responses from the data. One additional response was removed due to suspicion of close duplication. These lead us to our final data set of 255 valid responses.

After data cleaning, we tested the assumption of homogeneity by running Levene's test. Data transformation was performed (e.g., Normality transformations) if ANOVA assumptions were violated, and we reported results accordingly. We used a p-value $< .05$ to indicate statistical significance. Bonferroni corrections were performed for any post hoc comparisons as needed. We reported partial Eta-squared (η_p^2) for reporting effect size. Data were analyzed using SPSS.

For the qualitative data analysis, we used a bottom-up coding approach [54] to code the responses to the open-ended question. Initially, two researchers, r_1 and r_2, coded each statement independently by reading through all the comments. Then, both coders met and decided on the final codebook in the presence of a third researcher, r_3, who worked as a moderator and helped resolve the conflicts. Once the final codebook was generated, coders updated their version of the codebooks independently. Inter-rater reliability (IRR) for each question was calculated using Cohen's Kappa, which ranged from 0.7 to 0.97, indicating "substantial" or "excellent" agreement between the coders [51].

4 RESULTS

Among the 255 valid responses, there were 140 males (54.9%), 112 females (43.9%), one other, and 2 participants who chose not to answer. Participants' age ranged from 19 to 80 years ($M = 38.89, SD = 11.17$). Please refer to table 3 in the Appendix for a group-wise demographic breakdown of our participants.

A Kruskal-Wallis test found significant differences between the groups in terms of age ($H(7) = 19.849, p = .006$). Further post hoc test with Bonferroni correction showed significant differences between the groups 1A ($M = 35.15, SD = 9.5$) and 4B ($M = 43.75, SD = 11.23$). We ran Mann-Whitney tests for each independent variable and found a significant difference between user type and age ($U = 5725, p < .001$). No differences were found for account type ($U = 7564, p = .392$) and usage scenario ($U = 7637, p = .467$). From these analyses, we concluded that there were significant differences between adopters ($M = 36.23, SD = 10.25$) and non-adopters ($M = 41.6, SD = 11.45$) in terms of age.

A Chi-squared test revealed no significant differences between groups for gender ($\chi^2(21) = 23.453, p = .32$). A Kruskal-Wallis test reported no significant differences between groups in terms of the level of education ($H(7) = 10.201, p = .177$). We concluded that the groups are demographically comparable, except adopters were significantly younger than non-adopters, which is consistent with

Blame Attribution	Statement	Observed Difference
Blaming Bob	(Bob) Blaming himself for not taking the time to learn more.	Adopters attributed more blame than non-adopters towards Bob irrespective of account type and usage.
	(Bob) Blaming himself for not being more careful.	Adopters attributed more blame than non-adopters towards Bob when Bob used his email for intended purposes.
Blaming the Entity	Blaming the company for failing to communicate risk.	Non-adopter attributed more blame than adopters towards the company when Bob used his email for both official and personal purposes.
	Blaming the service provider for failing to communicate risk.	Non-adopters attributed more blame than adopters towards the service provider irrespective of account type and usage.
	Blaming the company for providing service with security vulnerabilities.	Non-adopters attributed more blame than adopters towards the company when Bob used his email for both official and personal purposes.
	Blaming the service provider for providing service with security vulnerabilities.	Non-adopters attributed more blame than adopters towards the service provider when Bob used his email for both official and personal purposes.
Blaming the Promotional Message	Advertisement message promoting 2-FA should have been clearer in explaining the purpose of two-step verification (2-FA).	Non-adopters agreed more than adopters when Bob used his email for both official and personal purposes.

Table 2: Key differences displayed by adopters and non-adopters at attributing blame towards different entities.

prior findings that noted that older adults are less likely to secure their device compared to younger adults [16].

To improve the paper's readability, we present a summary of the key differences between adopters and non-adopters in their blame attribution in Table 2, which we discuss in the following sections.

4.1 Users' Attribution of Blame towards Different Stakeholders

We asked participants to rate several statements related to their distribution of blame. The statements related to blame attribution were curated keeping in mind the different stakeholders associated with an attack and what possible reasons may be considered while blaming the stakeholders. We present our findings below.

4.1.1 Blaming the Attacker. When asked whether Bob should blame the attacker for the compromise, 78.1% (100/128) adopters and 74.8% (95/127) non-adopters somewhat to strongly agreed with the statement. ANOVA results showed no significant difference between the ratings given by the adopters ($M = 4.9, SD = 1.3$) and non-adopters ($M = 5.0, SD = 1.2$). However, we noticed a significant interaction effect between user and account type. Among the non-adopters, participants rated significantly lower when Bob's official email got compromised ($M = 4.8, SD = 1.3$), compared to when Bob's personal email got compromised ($M = 5.2, SD = 1.2$) ($F(1, 247) = 4.126, p = .035, \eta_p^2 = .016$). No such difference between official ($M = 4.88, SD = 1.2$) vs. personal ($M = 4.96, SD = 1.4$) email was apparent among the adopters.

4.1.2 Blaming Bob. We asked participants whether Bob should blame himself for not taking the responsibility to enable 2-FA for his email account. Among the adopters, 80.5% (103/128) somewhat to strongly agreed with the statement, with 58.6% agreeing strongly and 21.9% agreeing somewhat. Among the non-adopters, 78.7% (100/127) somewhat to strongly agreed, including 44.1% agreed strongly and 34.6% somewhat agreed. However, the ANOVA test showed no significant difference between adopters ($M = 5.2, SD = 1.2$) and non-adopters ($M = 5.06, SD = 1.2$).

Next, participants were asked to rate the statement, "Bob should blame himself for not taking the time to learn more about two-step verification (2-FA)." Participants mostly agreed with the statement indicating strong blame attribution towards Bob. 84.4% (108/128) of the adopters responded with somewhat or strong agreement compared to 72.4% (92/127) of the non-adopters. Unlike the first statement, ANOVA tests indicated a significant difference between adopters ($M = 5.3, SD = 1.1$) and non-adopters ($M = 4.9, SD = 1.3$) ($F(1, 247) = 5.7, p = .018, \eta_p^2 = .023$). We also noticed an interaction effect between usage scenario and user type. When Bob used his compromised email for both official and personal purposes, adopters ($M = 5.32, SD = 1.1$) blamed Bob more than non-adopters ($M = 4.9, SD = 1.3$). Although, the effect was marginally significant ($F(1, 247) = 3.839, p = .051, \eta_p^2 = .015$).

Lastly, in response to whether Bob should blame himself for not being more careful about the data present in his email, 83.6% (107/128) of the adopters and 72.4% (92/127) of the non-adopters agreed somewhat or strongly. This difference between adopters ($M = 5.374, SD = 1.0$) and non-adopters ($M = 4.938, SD = 1.2$) was significant ($F(1, 247) = 9.9, p = .002, \eta_p^2 = .038$). Similar to the second statement, we noticed a significant interaction effect between usage scenario and user type. Adopters ($M = 5.35, SD = 1.0$) seemed to blame Bob significantly more than non-adopters ($M = 4.81, SD = 1.2$) when Bob used his email for intended use only ($F(1, 247) = 7.6, p = .006, \eta_p^2 = .030$).

Graphs showing the differences between adopters and non-adopters in blaming Bob across the account and usage scenarios are presented in Appendix A.7.

In short, the above results indicate that adopters were likely to blame Bob more than non-adopters, specifically for not taking the time to learn more or for not being careful about the data in his email account. Thus, compared to the non-adopting population of our sample, adopters were more inclined to hold Bob responsible for his negligent behavior of not enabling 2-FA, which they believed led to the data breach (**Finding 1**).

4.1.3 Blaming the Entity. Interestingly, the direction of blame was reversed for the service provider or the company compared to Bob himself. 14.8% (19/128) of adopters and 22.8% (29/127) of non-adopters responded with somewhat or strong agreement when asked to attribute blame to the service provider for failing to communicate the risk adequately. ANOVA results indicated a significant main effect of user type. Non-adopters ($M = 3.2, SD = 1.4$) delegated the blame at a significantly higher rate than adopters ($M = 2.8, SD = 1.3$) towards the service provider regardless of account type and usage scenarios ($F(1, 247) = 5.3, p = .022, \eta_p^2 = .021$).

We also noticed differences across groups when asked participants whether Bob should blame the email service provider (personal email groups) for providing the service with security vulnerabilities. Among the participants, 21.1% (27/128) adopters rated somewhat or strongly agree compared to 32.3% (41/127) non-adopter participants. In addition, ANOVA results showed an interaction effect between account type and user type. For example, non-adopters ($M = 3.81, SD = 1.5$) blamed the service provider significantly more than adopters ($M = 3.2, SD = 1.4$) when Bob's personal email got compromised ($F(1, 247) = 5.95, p = .015, \eta_p^2 = .024$).

Another interaction effect between usage scenario and user type was also noticed. In particular, when Bob used his email for both official and personal purposes, adopters ($M = 2.7, SD = 1.3$) delegated less blame than non-adopters ($M = 3.6, SD = 1.4$). The difference was significant ($F(1, 247) = 12.27, p = .001, \eta_p^2 = .047$), suggesting that when Bob was using his email for multiple purposes, non-adopters blamed the service provider more for supposedly providing the service with security vulnerabilities.

Non-adopters also blamed the service provider more when Bob's personal email ($M = 3.81, SD = 1.5$) got compromised compared to when Bob's official email ($M = 3.27, SD = 1.4$) got compromised ($F(1, 247) = 4.6, p = .032, \eta_p^2 = .018$). Adopters did not show such distinction in blaming the service provider based on the account Bob was using. However, adopters did show significant differences in blaming the service provider for different usage scenarios. When Bob used his email for both official and personal purposes and got compromised, adopters attributed less blame ($M = 2.7, SD = 1.3$) to the provider compared to the case when Bob used his email for the intended purpose only ($M = 3.4, SD = 1.5$) ($F(1, 247) = 7.6, p = .006, \eta_p^2 = .03$).

Similarly, when asked whether Bob should blame the company (official email groups) for failing to communicate the risks adequately, 28.1% (36/128) adopters and 32.3% (41/127) non-adopters somewhat to strongly agreed with the statement. For the groups where Bob's official email got compromised, we noticed an interaction effect between usage scenario and user type. When Bob used his official email for both official and personal purposes and eventually got compromised, non-adopters ($M = 4.29, SD = 1.2$) blamed the company significantly more than adopters ($M = 3.4, SD = 1.5$) ($F(1, 124) = 5.4, p = .022, \eta_p^2 = .042$).

When inquired about blaming his company (official email groups) for providing services with security vulnerabilities, 37% (47/127) non-adopters and 31.3% (40/128) adopters agreed somewhat to strongly. Again, an interaction effect was noticed between the usage scenario and user type. Similar to the above statement, when Bob used his email for both official and personal purposes and eventually got compromised, non-adopters ($M = 3.8, SD = 1.4$) significantly blamed the company for providing service with security vulnerabilities than adopters ($M = 3.1, SD = 1.4$) ($F(1, 247) = 14.7, p = .01, \eta_p^2 = .027$).

Graphs showing the differences between adopters and non-adopters in blaming the company or the service provider across different account and usage scenarios are presented in Appendix A.8.

To summarize, our results showed that the adopters tended to blame the service provider or the company less than non-adopters. However, they blamed Bob more for his negligence than the non-adopters, hinting at the effect of "defensive attribution" [65] of blame in case of non-adopters (i.e., the tendency to blame external entities for an unintended incident) (**Finding 2**).

4.1.4 Attribution of Blame towards the Promotional Message. To understand participants' perception of the effectiveness of the promotional message (edited version of Google's official message), we asked participants whether the advertisement message promoting 2-FA should have been clearer in explaining the two-step verification (2-FA). Interestingly, 24.2% (31/128) adopters showed agreement compared to 29.1% (37/127) non-adopters, indicating that more non-adopters felt that the message was not clear enough in explaining what 2-FA is, which is also supported by the ANOVA results. We noticed a significant interaction effect between usage scenario and user type. When Bob used his email for both official and personal purposes, non-adopters ($M = 3.64, SD = 1.3$) agreed significantly more than the adopters ($M = 3.16, SD = 1.3$) about the message being less clear ($F(1, 247) = 4.25, p = .04, \eta_p^2 = .017$).

Alternatively, participants rated the message as less clear when Bob used his official email for only official purposes (intended use) ($M = 3.82, SD = 1.4$) compared to when he used his personal email for only personal purposes ($M = 3.13, SD = 1.2$) ($F(1, 247) = 9.2, p = .003, \eta_p^2 = .036$). The finding suggests that, when Bob used his personal email for personal purposes only, participants did not blame the message as much as when Bob used his official email for official purposes only.

Overall, non-adopters were more likely to rate the message as less effective, and felt that the message failed to explain 2-FA comprehensively. This further underscores the tendency to blame external factors among the non-adopters (**Finding 3**).

4.1.5 Attribution of Blame to Secondary Parties. Although not directly responsible for the attack, blame can be attributed to parties that may indirectly influence the negligent behavior displayed by users (e.g., Bob). We explored whether participants attribute any blame toward three of such parties - Bob's friends, colleagues, and the government. Participants did not blame them much for the attack. For example, when asked whether Bob should blame his friends for not informing him about 2-FA, 11% (28/255) of participants somewhat or strongly agreed. Similarly, when participants were told to attribute blame to Bob's colleagues for the same reason, only 10.2% (26/255) of participants agreed. When asked whether Bob should blame the government for not raising public awareness regarding the importance of 2-FA, 11% (28/255) participants indicated agreement. No significant differences were found among the groups.

4.2 Justification for Non-adopting Behavior and Defensive Attribution of Blame

To better understand the reasonings behind the attribution of blame, we asked participants about different aspects of the attack, such as whether the attack was realistic to happen, whether not enabling 2-FA was the right decision made by Bob, and whether enabling 2-FA would have prevented the attack or not. Our findings are presented below.

4.2.1 How likely is the attack mentioned in the vignettes? After part 2 of the vignette, participants were asked whether launching an attack similar to the one presented in the story was possible. Towards that, 92.2% (118/128) of the adopters responded yes to the question compared to 81.1% (103/127) of the non-adopters. A Chi-squared test indicated significant differences between adopters and non-adopters ($\chi^2(2) = 7.386, p = .025$). However, further investigation showed no significant differences between official or personal email groups ($\chi^2(2) = 2.304, p = .316$) and intended vs. combined usage groups ($\chi^2(2) = 0.005, p = .997$).

When asked about the likelihood of someone experiencing an attack of a similar nature, 73% (94/128) adopters rated somewhat to extremely likely, compared to 65% non-adopters. However, a Chi-squared test showed no significant differences between the two groups ($\chi^2(1) = 1.962, p = .161$). Furthermore, no significant differences between the groups were noticed for account type ($\chi^2(1) = 0.098, p = .754$) or usage scenario ($\chi^2(2) = .176, p = .675$).

Low perception regarding the possibility of launching a similar attack may cause non-adopters to view the attack as just "bad luck", making them feel that Bob was less responsible for the breach, thereby prompting them to delegate blame towards the externals. Hence, the findings indicate a gap in perception regarding the technical feasibility of the attack between adopters and non-adopters, whereas there is no significant gap in the likelihood of the attack (**Finding 4**).

4.2.2 Did Bob do the right thing by not enabling 2-FA?. We asked participants to rate the statement "Bob did the right thing by not enabling two-step verification (2-FA)". Overall, 70.2%(179/255) participants either strongly or somewhat disagreed with the statement, suggesting that the majority thought Bob did not do the right thing by completely ignoring the message. We found a significant main effect of user type for this statement. ANOVA results showed significant differences between adopters ($M = 1.837, SD = 1.2$) and non-adopters ($M = 2.3, SD = 1.1$), indicating that adopters significantly disagreed more with the statement than non-adopters irrespective of account type and usage scenario ($F(1, 247) = 10.3, p = .002, \eta_p^2 = .04$). Hence, adopters disagreed significantly more with Bob's decision not to enable 2-FA than non-adopters (**Finding 5**).

In addition to the quantitative results, we asked participants to mention their reasons for agreeing or disagreeing with whether Bob did the right thing by not enabling 2-FA. Table 4 in the Appendix shows the codes for this statement reflecting the reasoning behind both agreement and disagreement. We discuss the findings below. *Note that the comments presented in the paper are fixed for typos without altering the meaning.*

To explain why Bob should have enabled 2-FA, 46% (117/255) of the participants indicated the benefits of 2-FA. Among them, 57% (73/128) adopters and 35% (44/127) non-adopters mentioned how 2-FA could be beneficial for accounts. Participants mainly indicated the security benefits of the tool, with phrases like *extra security* or *added layer of security*. Some participants, primarily adopters (83%, 5/6), went further and stated that enabling 2-FA can give *peace of mind*.

27% (69/255) participants pointed out the *sensitive nature of the data* and/or Bob's *obligation towards the company* as important reasons for disagreeing with his decision. 75% (52/69) of these comments are from the official email group participants.

> *"Bob was wrong to not enable 2-FA for his office email account because the emails usually contain many sensitive customer data. Once the email password was stolen by bad guys, the customer data would be compromised." (GR 1B)*

Comments also went in the other direction and attempted to explain the reasoning behind Bob's decision not to enable 2-FA. For example, the inconvenience of using 2-FA was mentioned by 7% (18/255) of participants. 76% (13/17) of those comments came from the personal email group participants.

> *"You might need to get into your email and not have your phone or maybe you forget your password and don't want to take the extra steps to wait for a text and change your password." (GR 3A)*

6% (16/255) participants argued about how it should be Bob's *personal choice* to enable the feature for his account, especially when it was his personal account (81%, 13/16).

> *"It is his own choice and I applaud anyone for making decisions about their personal space." (GR 4A)*

4.2.3 Do Users Believe in 2-FA?. Regarding the efficacy of 2-FA, participants were asked whether enabling 2-FA would have prevented the attack mentioned in the vignette. Results showed that 54.7% (70/128) adopters and 48% (61/127) non-adopters expressed strong or somewhat disagreement with the statement "Enabling two-step verification (2-FA) would not have prevented the attack." However, ANOVA showed no significant differences among the groups.

Note that the term *attack* could refer to one of the two compromises mentioned in the story, namely, (a) compromise of the authentication server and (b) compromise of Bob's email account, which is reflected in the following comment.

> *".... I believe that the attack itself would not have been prevented (leak of passwords) but the repercussions of the attack (accessed accounts, changed passwords) could have been prevented with 2-FA." (GR 1A)*

To account for this dual interpretation, we excluded comments that were unclear about the compromise, which led us to drop 15% of the comments while coding. These comments were coded as "vague/unable to code"(Table 5 in the Appendix). Nonetheless, this item led many participants to explicitly think of and discuss the two different compromises, which we found valuable for understanding users' perception of the underlying 2-FA mechanism.

Qualitative coding and analysis showed that adopters and non-adopters did not significantly differ in explaining how 2-FA would have been effective for protecting Bob's email account. Overall,

61% (78/128) comments from the adopters and 51% (65/127) comments from the non-adopters indicated this fact. Adopters and non-adopters were surprisingly close in understanding how an additional step would have been required for the attackers to break in if Bob had enabled 2-FA for his email. 27% (35/128) comments from the adopters and 21% (27/127) comments from the non-adopters across the groups had explicitly mentioned the requirement of an extra step via a secondary device (e.g., a cellphone), suggesting a clear understanding of how the 2-FA mechanism works, as reflected in the following comment from a non-adopter participant.

> *"It would have likely prevented it as having a phone or trusted device would usually be needed in order to verify the account and the hacker wouldn't have had one." (GR 4B)*

Moreover, adopters and non-adopters were surprisingly similar in pointing out how the server attack would not have been prevented even if Bob had enabled 2-FA. For example, 13% (17/128) comments from the adopters and 9% (11/127) comments from the non-adopters across the groups mentioned this fact explicitly, as shown in the following non-adopter comment.

> *"The attack itself wouldn't have been prevented by 2-FA, I don't think, because they attacked a database or something and got lots of login credentials...." (GR 1B)*

Interestingly, adopters were closely in line with non-adopters in showing skepticism about the effectiveness of 2-FA. Overall, about 30% (73/255) of the participant comments indicated such skepticism, distributed by 27% (34/128) adopter comments and 31% (39/127) non-adopter comments. Displayed skepticism was not always attributed to a poor understanding of 2-FA; instead, it was quite the opposite. Even participants who understood what 2-FA is and how the underlying mechanism works showed skepticism about the effectiveness of 2-FA.

> *"It would be very difficult for the attackers to have access to the cell phone or device used for the second verification. They likely would not be able to access his email. However, they broke into the server, so I don't know enough about IT to know if they possibly could have accessed it anyway." (GR 3B)*

A theme of *"2-FA makes it harder but may not stop"* was prominent across participants' comments. Overestimating the ability of the attackers was a prominent reason for skepticism, as displayed in the following comment from an adopter.

> *"I think enabling the two-step verification would have made it harder for the hacker to steal his login credentials. But at the same time, if the hacker is really advanced they might be able to bypass any type of verification system." (GR 1A)*

Overall, the analysis identified two key observations. First, non-adopters had a similar mental model as adopters in understanding what 2-FA is, how the feature works, and how it increases security, even though they chose not to adopt it. This observation is in line with what Herley argued, namely, end-users understand risks correctly and make a conscious decision to follow or not follow a security practice/recommendation [40]. Second, even though most participants agreed that 2-FA increases the security of the account and makes it harder for the attackers to break, about one-third of the participants in both groups were erroneously hesitant to declare that enabling 2-FA would have kept Bob secure, which shows a low perceived response efficacy of 2-FA in the mental models of both adopters and non-adopters (**Finding 6**).

5 DISCUSSION

Our study showed that the adopters and non-adopters are, in fact, fundamentally different in the way they attribute blame after a data breach. The implications of our findings are presented below. A summary of the key implications for the pre-breach and post-breach communication is presented in Table 6 in the Appendix.

5.1 Implications for Pre and Post-Security Breach Risk Communication

In our study, adopters indicate a tendency of "self-blame" compared to the tendency of "other-blame" among the non-adopters. However, it is not clear what causes this difference. One possibility is that these two groups of users (i.e., adopters vs. non-adopters) are fundamentally different in personality traits and perception of self-responsibility, and these divergences in personality lead to differences in behavior. On the other hand, some may argue that behavior change comes first, which subsequently leads to differences in perceptions regarding personal responsibility. This argument can be supported using the theory of cognitive dissonance, which suggests that once users adopt a behavior (either voluntarily or due to a requirement), users' reasoning behind the behavior aligns with their behavior to minimize cognitive discomfort [30]. Irrespective of what drives the sense of self-responsibility and causes this difference, attribution of self-blame is considered a spontaneous affective and cognitive self-evaluative reaction [8]. Importantly, social cognitive frameworks showed evidence that attribution of self-blame successfully predicts positive coping response behavior, whereas blaming others predicts poor coping response [17, 56, 66]. Further, work in psychology and healthcare showed that attributing blame to oneself creates a belief of self-responsibility, which leads to proactive behavior targeting improvement [44, 56]. Thus, the tendency of "self-blame" among the adopters indicates that they are more likely to exhibit positive coping response behavior and follow the recommended actions after a security breach (e.g., enabling credit monitoring, changing password) compared to the non-adopting population who tend to delegate blame to others.

On the other hand, non-adopting participants in our study were more likely to hold the external entities responsible. Our results align with prior efforts showing that personal similarity with the stimulus actor can induce defensive attribution [64, 65]. Hence, it is likely that, due to the perception of similarity to Bob (non-adopter of 2-FA), this participant group indicated more lenience towards Bob and (defensively) blamed him less. Prior work noted the negative impact of defensive attribution, such as avoidance of responsibility, overreliance on external factors, and negative coping response [12, 18, 33]. Thus, following a data breach, the non-adopting population is less likely to change their behavior to protect themselves by adopting security measures. A similar observation was reported in Zou et al.'s work after the Equifax data breach, where, despite being concerned, most participants did not take any post-breach measure [82]. Furthermore, with the belief

that service providers are responsible for security, a data breach incident is likely to cause feelings of breach of trust substantially more among the non-adopting group, requiring companies to take part in extensive trust repair effort following a breach incident [9, 52]. On the other hand, adopters are less likely to experience psychological trust violations following a data breach incident. Thus, post-breach communication strategies should be careful about incorporating the aspect of self-responsibility to calibrate the cognitive and affective processes that can influence the decision of the non-adopting population. If not done right, promoting personal responsibility after a data breach may be perceived as an attempt to shift blame to users and can trigger further resistance to change.

Our research also has implications for pre-breach risk communication. Specifically, as we noticed that non-technical participants are more likely to shift responsibility after a breach, pre-breach messaging should underscore the importance of shared responsibility that can change the perception of self-responsibility and promote concerted team effort where both parties acknowledge shared responsibilities and act responsibly.

5.2 Mindful Consideration of Context

Security tools/behavior promotional messages are currently designed with a "one-size-fits-all" mindset, meaning the same message is delivered to millions of users without considering the context. However, based on our findings, usage contexts influence users' perceived effectiveness of the promotional message, which can influence their decisions. For instance, participants seemed to assign more responsibility to the end-user (e.g., Bob) when the official email was in context. However, this observation also indicates a misconception of low perceived data value regarding personal accounts, which is noted in prior work as well [28]. As low perceived data value can potentially dampen the importance of protecting the account, communication of data value is essential. Interestingly, although the same promotional message was used across all the groups, the comprehensibility ratings differed based on the account and usage contexts, which can indicate non-adopting users' tendency to shift blame to external factors, especially when the consequence is high. Again, this is consistent with the findings of Shaver et al. [65], where non-technical participants were found defensively shifting their attribution of responsibility to others.

These findings suggest the importance of considering contexts while designing promotional messages for communicating risks in pre and post-breach scenarios. Specifically, users are more likely to pay closer attention to security recommendations when they feel the need to protect their accounts. This is likely to happen if the communications raise awareness regarding perceived data values and the cost of compromise. As such, context-sensitive promotional messages for promoting security tools/behavior that consider users' specific usage patterns have a higher chance of success. Towards that, one possibility is for companies and service providers to take approaches similar to Ad personalization for delivering personalized promotional messages for security tools based on usage data.

5.3 Misunderstanding Surrounding the Response Efficacy of Security Tools

A surprising finding in our study was the degree of skepticism among both groups regarding the efficacy of 2-FA in preventing similar attacks. In particular, participants believed that, even though 2-FA increases the security of the accounts and makes it harder for the attackers to breach successfully, it is not fail-proof. However, it is not evident whether the skepticism comes from their poor evaluation of 2-FA or the overestimation of the ability of the attackers to break down almost anything. Overestimation of attackers' ability, which most likely comes from their misconceptions around security and/or media portrayal [31], has been shown in prior efforts [35, 75].

Unfortunately, in pre-security breach communication, if this skepticism is not addressed, this can lead to a sense of "hopelessness" and play a role in their non-compliant behavior. This can play an even more prominent role in post-breach communication, as users are likely to feel betrayed by the breach and ignore further security advice due to a lack of perceived response efficacy, which is noted as an important antecedent of behavior change in a number of theoretical frameworks (e.g., Protection Motivation Theory (PMT)). As such, in addition to explaining the benefit, it is crucial to explicitly communicate under what circumstances a recommended security feature will not work and what the responsibilities are on the users' end for the feature to work securely. Notably, given that non-technical participants mostly hold the service provider responsible, after a security breach, any communication should explain why the existing security features failed to prevent the attack and how the recommended action will minimize the risk in the future.

5.4 Addressing Low Perceived Vulnerability

Although numerous cyberattacks have been publicized in the media in recent times [10, 15, 43, 46, 61], however, non-adopters found the attack to be less likely to happen in real life compared to the adopters, indicating a low perceived vulnerability related to cyberattacks. This finding can be explained by "optimistic bias" [78] and is in line with prior efforts that demonstrated similar phenomena in other contexts such as screen locking behavior [5] and underestimating the value of data, which can contribute to a lack of concern [3, 60, 82]. This finding underscores the importance of information campaigns targeting users' low perceived vulnerability and incorporating such messaging while promoting cybersecurity tools/behavior, both in the context of pre and post-breach communication.

5.5 Limitations of the Study

Our study identifies important differences between adopting and non-adopting populations regarding attribution of blame. However, our findings should be interpreted with the following limitations in mind.

First, we restricted our participants to the adult population who currently live in the United States and have the technical ability to use MTurk, which may not be a representative population of the United States. Further, the blame attribution observed in the context of 2-FA could be different compared to other security tools usage scenarios (e.g., antivirus software, password managers), which is essential to understanding our findings' generalizability.

Second, we used Bob as the protagonist instead of asking participants to imagine themselves in place of Bob. It was done to avoid possible "self-serving bias" [11] that can nudge participants

to blame others for deflecting blame. Further, we did not ask participants whether they felt similar to Bob to avoid triggering social desirability bias [27], which would have defeated the purpose of using vignettes in the first place. The study aimed to investigate blame attribution from a neutral perspective, which can facilitate further investigation and design of persuasive messages taking blame attribution factors into account. Interestingly, participants' responses in our study suggested that, even though people tend to blame others, they know, at least to a certain degree, that they are responsible for their actions. Further studies are required to explore and compare blame attribution in real life vs. hypothetical scenarios in the context of a security breach.

Third, as our exploratory study aimed to investigate divergences in blame attribution based on past compliance behavior, we did not explicitly control for age or other demographic variables. The observed age difference between adopters and non-adopters is consistent with prior findings that noted that older adults are less likely to secure their devices compared to younger adults [16]. Nonetheless, age (and other demographic factors) could be associated with certain aspects of blame attribution. As such, based on our findings, future efforts can design studies considering age as an independent variable and look into divergences in blame attribution across different age groups. Our results should be interpreted keeping the observed group differences in terms of the demographic factor (i.e., age) in mind.

Finally, in a vignette-based study like ours, certain attribution biases (e.g., fundamental attribution error [63], actor-observer bias [47]) can occur and influence participants' responses. However, we chose a vignette-based study as it offers several advantages over a direct question-based study, such as offering greater realism, allowing delivery of standard stimuli to all respondents (which enhances internal validity), and reducing social desirability bias [27]. To minimize any potential bias, in our case, the vignette wordings were the same for both the adopting participants and the non-adopting participants except for the account types and usage scenarios as necessary, thereby ensuring that the differences observed across the groups were due to the treatments used in the study.

6 CONCLUSION

In this work, we explore how participants distribute blame toward different parties that may be held responsible for a data breach. Findings indicate that adopters of a tool are more likely to hold the end-users responsible for their negligent behavior even when they are the victims. On the other hand, non-adopters show defensive attribution of blaming and hold the service providers more responsible. Our findings confirm the gap in responsibility perception between the two groups. This gap in perception provides support in favor of promoting personal responsibility as a viable component in risk communication messages while promoting security tools/behaviors, both in pre and post-breach contexts. Further, our results show that attribution of blame is context-dependent and may change based on account type and usage scenarios. As such, future research should explore the possibility of personalizing intervention messages and test their efficacy in different contexts instead of designing "one-size-fits-all" messages.

ACKNOWLEDGMENTS

This research was supported by a NSF CAREER award to the second author, 1750908.

REFERENCES

[1] Ezaz Ahmed, ABM Abdullah, and Md Wahid Murad. 2020. Relationship between Psychological Contract Breach and Employee Outcomes: Moderating Role of Blame Attribution. *South Asian Journal of Management* 27, 3 (2020).
[2] Icek Ajzen. 1991. The theory of planned behavior. *Organizational behavior and human decision processes* 50, 2 (1991), 179–211.
[3] Mahdi Nasrullah Al-Ameen and Huzeyfe Kocabas. 2020. "I cannot do anything": User's Behavior and Protection Strategy upon Losing, or Identifying Unauthorized Access to Online Account. In *Symposium on Usable Privacy and Security (Poster Session)*.
[4] Yusuf Albayram, Mohammad Maifi Hasan Khan, and Michael Fagan. 2017. A study on designing video tutorials for promoting security features: A case study in the context of two-factor authentication (2FA). *International Journal of Human–Computer Interaction* 33, 11 (2017), 927–942.
[5] Yusuf Albayram, Mohammad Maifi Hasan Khan, Theodore Jensen, and Nhan Nguyen. 2017. "... better to use a lock screen than to worry about saving a few seconds of time": Effect of Fear Appeal in the Context of Smartphone Locking Behavior. In *Thirteenth Symposium on Usable Privacy and Security ({SOUPS} 2017)*. 49–63.
[6] Farzaneh Asgharpour, Debin Liu, and L Jean Camp. 2007. Mental Models of Computer Security Risks.. In *WEIS*.
[7] Emmanuel W Ayaburi and Daniel N Treku. 2020. Effect of penitence on social media trust and privacy concerns: The case of Facebook. *International Journal of Information Management* 50 (2020), 171–181.
[8] Albert Bandura. 1991. Social cognitive theory of self-regulation. *Organizational behavior and human decision processes* 50, 2 (1991), 248–287.
[9] Gaurav Bansal and Noah Redfearn. 2019. Trust Violation and Rebuilding After a Data Breach: Role of Environmental Stewardship and Underlying Motives. *Journal of the Midwest Association for Information Systems (JMWAIS)* 2019, 2 (2019), 4.
[10] Gabriel Bassett, C David Hylender, Philippe Langlois, Alexandre Pinto, and Suzanne Widup. 2021. Data breach investigations report. *Verizon DBIR Team, Tech. Rep* (2021).
[11] RR Baumeister. 1998. The self (In DT Gilbert, ST Fiske, & G. Lindzey (Eds.). The handbook of social psychology (Vol. 1, pp. 680–740). *NY: McGraw-Hill* (1998).
[12] Roy F Baumeister and Steven J Scher. 1988. Self-defeating behavior patterns among normal individuals: review and analysis of common self-destructive tendencies. *Psychological bulletin* 104, 1 (1988), 3.
[13] Zinaida Benenson, Freya Gassmann, and Lena Reinfelder. 2013. Android and iOS users' differences concerning security and privacy. In *CHI'13 Extended Abstracts on Human Factors in Computing Systems*. 817–822.
[14] Sruti Bhagavatula, Lujo Bauer, and Apu Kapadia. 2020. What breach? Measuring online awareness of security incidents by studying real-world browsing behavior. *arXiv preprint arXiv:2010.09843* (2020).
[15] Paul Bischoff. 2019. 7 million Adobe Creative Cloud accounts exposed to the public.
[16] Dawn Branley-Bell, Lynne Coventry, Matt Dixon, Adam Joinson, Pam Briggs, et al. 2022. Exploring age and gender differences in ICT cybersecurity behaviour. *Human Behavior and Emerging Technologies* 2022 (2022).
[17] Ronnie J Bulman and Camille B Wortman. 1977. Attributions of blame and coping in the" real world": severe accident victims react to their lot. *Journal of personality and social psychology* 35, 5 (1977), 351.
[18] Jerry M Burger. 1981. Motivational biases in the attribution of responsibility for an accident: A meta-analysis of the defensive-attribution hypothesis. *Psychological Bulletin* 90, 3 (1981), 496.
[19] Karoline Busse, Julia Schäfer, and Matthew Smith. 2019. Replication: no one can hack my mind revisiting a study on expert and non-expert security practices and advice. In *Fifteenth Symposium on Usable Privacy and Security ({SOUPS} 2019)*.
[20] Christopher J Carpenter. 2010. A meta-analysis of the effectiveness of health belief model variables in predicting behavior. *Health communication* 25, 8 (2010), 661–669.
[21] Victoria L Champion and Celette Sugg Skinner. 2008. The health belief model. *Health behavior and health education: Theory, research, and practice* 4 (2008), 45–65.
[22] Sandra Costa and Pedro Neves. 2017. It is your fault! How blame attributions of breach predict employees' reactions. *Journal of Managerial Psychology* (2017).
[23] Sanchari Das, Jacob Abbott, Shakthidhar Gopavaram, Jim Blythe, and L Jean Camp. 2020. User-Centered Risk Communication for Safer Browsing. In *Proceedings of the First Asia USEC-Workshop on Usable Security, In Conjunction with the Twenty-Fourth International Conference International Conference on Financial Cryptography and Data Security*.

[24] Sidney Dekker. 2013. *Second victim: error, guilt, trauma, and resilience.* CRC press.

[25] Sidney Dekker. 2018. *Just culture: restoring trust and accountability in your organization.* CRC press.

[26] Dennis J Devine and David E Caughlin. 2014. Do they matter? A meta-analytic investigation of individual characteristics and guilt judgments. *Psychology, Public Policy, and Law* 20, 2 (2014), 109.

[27] Allen L Edwards. 1957. The social desirability variable in personality assessment and research. (1957).

[28] Serge Egelman, Sakshi Jain, Rebecca S Portnoff, Kerwell Liao, Sunny Consolvo, and David Wagner. 2014. Are you ready to lock?. In *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security*. 750–761.

[29] Michael Fagan and Mohammad Maifi Hasan Khan. 2016. Why do they do what they do?: A study of what motivates users to (not) follow computer security advice. In *Twelfth Symposium on Usable Privacy and Security ({SOUPS} 2016)*. 59–75.

[30] Leon Festinger. 1957. *A theory of cognitive dissonance*. Vol. 2. Stanford university press.

[31] Kelsey R Fulton, Rebecca Gelles, Alexandra McKay, Yasmin Abdi, Richard Roberts, and Michelle L Mazurek. 2019. The effect of entertainment media on mental models of computer security. In *Fifteenth Symposium on Usable Privacy and Security ({SOUPS} 2019)*.

[32] Gemalto. 2016. Data Breaches and Customer Loyalty 2016. *Technical Report* (2016).

[33] Jeff Greenberg, Tom Pyszczynski, and Sheldon Solomon. 1982. The self-serving attributional bias: Beyond self-presentation. *Journal of Experimental Social Psychology* 18, 1 (1982), 56–67.

[34] M. Hadi and B. Logan. 2017. Equifax: Hackers may have the personal details of 143 million US customers. http://www.businessinsider.com/equifaxhackers-may-have-accessed-personal-details143-million-us-customers-2017-9. last accessed on: 01.22.2018.

[35] Julie M Haney and Wayne G Lutters. 2018. " It's Scary... It's Confusing... It's Dull": How Cybersecurity Advocates Overcome Negative Perceptions of Security. In *Fourteenth Symposium on Usable Privacy and Security ({SOUPS} 2018)*. 411–425.

[36] Marian Harbach, Markus Hettig, Susanne Weber, and Matthew Smith. 2014. Using personal examples to improve risk communication for security & privacy decisions. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 2647–2656.

[37] Marian Harbach, Emanuel Von Zezschwitz, Andreas Fichtner, Alexander De Luca, and Matthew Smith. 2014. It'sa hard lock life: A field study of smartphone (un) locking behavior and risk perception. In *10th Symposium On Usable Privacy and Security ({SOUPS} 2014)*. 213–230.

[38] Paul Harvey and Marie T Dasborough. 2006. Consequences of employee attributions in the workplace: The role of emotional intelligence. *Psicothema* (2006), 145–151.

[39] Fritz Heider. 2013. *The psychology of interpersonal relations*. Psychology Press.

[40] Cormac Herley. 2009. 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*. 133–144.

[41] Adele E Howe, Indrajit Ray, Mark Roberts, Malgorzata Urbanska, and Zinta Byrne. 2012. The psychology of security for the home computer user. In *2012 IEEE Symposium on Security and Privacy*. IEEE, 209–223.

[42] Iulia Ion, Rob Reeder, and Sunny Consolvo. 2015. "... no one can hack my mind": Comparing Expert and Non-Expert Security Practices. In *Eleventh Symposium On Usable Privacy and Security ({SOUPS} 2015)*. 327–346.

[43] Mike Isaac and Sheera Frenkel. 2018. Facebook Security Breach Exposes Accounts of 50 Million Users (Published 2018)" – NYTimes.com.

[44] Ronnie Janoff-Bulman. 1979. Characterological versus behavioral self-blame: Inquiries into depression and rape. *Journal of personality and social psychology* 37, 10 (1979), 1798.

[45] Ronnie Janoff-Bulman. 1982. Esteem and control bases of blame:"Adaptive" strategies for victims versus observers. *Journal of personality* 50, 2 (1982), 180–192.

[46] Martin Jartelius. 2020. The 2020 Data Breach Investigations Report–a CSO's perspective. *Network Security* 2020, 7 (2020), 9–12.

[47] Edward E Jones and Richard E Nisbett. 1987. The actor and the observer: Divergent perceptions of the causes of behavior.. In *Preparation of this paper grew out of a workshop on attribution theory held at University of California, Los Angeles, Aug 1969*. Lawrence Erlbaum Associates, Inc.

[48] Daniel Kahneman and Amos Tversky. 2013. Choices, values, and frames. In *Handbook of the fundamentals of financial decision making: Part I*. World Scientific, 269–278.

[49] Ruogu Kang, Laura Dabbish, Nathaniel Fruchter, and Sara Kiesler. 2015. "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)*. 39–52.

[50] Shipi Kankane, Carlina DiRusso, and Christen Buckley. 2018. Can We Nudge Users Toward Better Password Management? An Initial Study. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*. 1–6.

[51] J Richard Landis and Gary G Koch. 1977. The measurement of observer agreement for categorical data. *biometrics* (1977), 159–174.

[52] Kristin Masuch, Maike Greve, and Simon Trang. 2021. What to do after a data breach? Examining apology and compensation as response strategies for health service providers. *Electronic Markets* 31 (2021), 829–848.

[53] Miriam J Metzger. 2006. Effects of site, vendor, and consumer characteristics on web site trust and disclosure. *communication research* 33, 3 (2006), 155–179.

[54] Matthew B Miles and A Michael Huberman. 1994. *Qualitative data analysis: An expanded sourcebook*. sage.

[55] Elizabeth Wolfe Morrison and Sandra L Robinson. 1997. When employees feel betrayed: A model of how psychological contract violation develops. *Academy of management Review* 22, 1 (1997), 226–256.

[56] Jeffrey M Moulton, David M Sweet, Lydia Temoshok, and Jeffrey S Mandel. 1987. Attributions of Blame and Responsibility in Relation to Distress and Health Behavior Change in People with AIDS and AIDS-related Complex 1. *Journal of Applied Social Psychology* 17, 5 (1987), 493–506.

[57] Sarah Peck, Mohammad Khan, Md Fahim, Emil Coman, Theodore Jensen, and Yusuf Albayram. 2020. Who Would Bob Blame? Factors in Blame Attribution in Cyberattacks Among the Non-adopting Population in the Context of 2FA.. In *IEEE 44th Annual Computers, Software, and Applications Conference (COMPSAC)*.

[58] Ahmad R Pratama and Firman M Firmansyah. 2021. Until you have something to lose! Loss aversion and two-factor authentication adoption. *Applied Computing and Informatics* (2021).

[59] Fahimeh Raja, Kirstie Hawkey, Steven Hsu, Kai-Le Wang, and Konstantin Beznosov. 2011. Promoting a physical security mental model for personal firewall warnings. In *CHI'11 Extended Abstracts on Human Factors in Computing Systems*. ACM, 1585–1590.

[60] Karen Renaud, Melanie Volkamer, and Arne Renkema-Padmos. 2014. Why doesn't Jane protect her privacy?. In *International Symposium on Privacy Enhancing Technologies Symposium*. Springer, 244–262.

[61] The Published Reporter. 2019. "Nearly 12,000,000 Quest Diagnostics Patients' Medical Info Exposed In New Data Breach Of Third-Party Billing Collections Vendor". The Published Reporter.

[62] Ronald W Rogers and Steven Prentice-Dunn. 1997. Protection motivation theory. *Handbook of health behavior research 1: Personal and social determinants* (1997), 113–132.

[63] Lee Ross. 1977. The intuitive psychologist and his shortcomings: Distortions in the attribution process. In *Advances in experimental social psychology*. Vol. 10. Elsevier, 173–220.

[64] Nir Rozmann and Inna Levy. 2021. Attribution of blame toward offenders: Victim and offender ethnicity, and observer ethnic and religious background. *Journal of interpersonal violence* 36, 21-22 (2021), 10638–10659.

[65] Kelly G Shaver. 1970. Defensive attribution: Effects of severity and relevance on the responsibility assigned for an accident. *Journal of personality and social psychology* 14, 2 (1970), 101.

[66] Diane E Sholomskas, Janice M Steil, and Jack K Plummer. 1990. The Spinal Cord Injured Revisited: The Relationship Between Self-Blame, Other-Blame and Coping. *Journal of Applied Social Psychology* 20, 7 (1990), 548–574.

[67] Michael Warren Skirpan, Tom Yeh, and Casey Fiesler. 2018. What's at Stake: Characterizing Risk Perceptions of Emerging Technologies. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 1–12.

[68] Artur Strzelecki and Mariia Rizun. 2022. Consumers' Change in Trust and Security after a Personal Data Breach in Online Shopping. *Sustainability* 14, 10 (2022), 5866.

[69] Henri Tajfel and John C Turner. 1982. Social psychology of intergroup relations. *Annual review of psychology* 33, 1 (1982), 1–39.

[70] Hsin-yi Sandy Tsai, Mengtian Jiang, Saleem Alhabash, Robert LaRose, Nora J Rifon, and Shelia R Cotten. 2016. Understanding online safety behaviors: A protection motivation theory perspective. *Computers & Security* 59 (2016), 138–150.

[71] Hilary Tuttle. 2017. The data breach blame game. *Risk Management* 64, 2 (2017), 36–37.

[72] Ehsan Ul Haque, Mohammad Maifi Hasan Khan, and Md Abdullah Al Fahim. 2023. The Nuanced Nature of Trust and Privacy Control Adoption in the Context of Google. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–23.

[73] Blase Ur, Fumiko Noma, Jonathan Bees, Sean M Segreti, Richard Shay, Lujo Bauer, Nicolas Christin, and Lorrie Faith Cranor. 2015. " I Added'!'at the End to Make It Secure": Observing Password Creation in the Lab. In *Eleventh Symposium On Usable Privacy and Security ({SOUPS} 2015)*. 123–140.

[74] Kami E Vaniea, Emilee Rader, and Rick Wash. 2014. Betrayed by updates: how negative experiences affect future security. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 2671–2674.

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

[76] Rick Wash and Emilee Rader. 2015. Too much knowledge? security beliefs and protective behaviors among united states internet users. In *Eleventh Symposium On Usable Privacy and Security ({SOUPS} 2015)*. 309–325.
[77] Bernard Weiner. 2012. *An attributional theory of motivation and emotion*. Springer Science & Business Media.
[78] Neil D Weinstein and William M Klein. 1996. Unrealistic optimism: Present and future. *Journal of Social and Clinical Psychology* 15, 1 (1996), 1–8.
[79] Charles Cresson Wood and William W Banks Jr. 1993. Human error: an overlooked but significant information security problem. *Computers & Security* 12, 1 (1993), 51–60.
[80] Allison Woodruff, Vasyl Pihur, Sunny Consolvo, Laura Brandimarte, and Alessandro Acquisti. 2014. Would a Privacy Fundamentalist Sell Their {DNA} for $1000... If Nothing Bad Happened as a Result? The Westin Categories, Behavioral Intentions, and Consequences. In *10th Symposium On Usable Privacy and Security ({SOUPS} 2014)*. 1–18.
[81] Yixin Zou, Shawn Danino, Kaiwen Sun, and Florian Schaub. 2019. YouMight'Be Affected: An Empirical Analysis of Readability and Usability Issues in Data Breach Notifications. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–14.
[82] Yixin Zou, Abraham H Mhaidli, Austin McCall, and Florian Schaub. 2018. " I've Got Nothing to Lose": Consumers' Risk Perceptions and Protective Actions after the Equifax Data Breach. In *Fourteenth Symposium on Usable Privacy and Security ({SOUPS} 2018)*. 197–216.
[83] Yixin Zou and Florian Schaub. 2018. Concern But No Action: Consumers' Reactions to the Equifax Data Breach. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*. 1–6.

A APPENDIX

A.1 Survey Instrument

A.1.1 Prescreening Questionnaire.

(1) Are you proficient in English?
- Definitely yes
- Probably yes
- Might or might not
- Probably not
- Definitely not

(2) Do you have a degree in Computer Science, including a "minor," or any professional computer science certifications?
- Yes
- No

(3) Do you use email for important tasks (e.g., official business, receiving notifications from bank or other financial institution, online bill payment, other important online services etc.)?
- Yes
- No

(4) Which of the following email account(s) do you use for important tasks (e.g., official business, receiving notifications from bank or other financial institution, online bill payment, other important online services etc.)? Please select all that apply.
- Gmail
- Outlook
- Yahoo Main
- Others (Please specify)

(5) Do you know what two-factor authentication (a.k.a. 2-FA or two-step verification) is?
- Yes
- No

(6) Do you currently use two-factor authentication (a.k.a. 2-FA or two-step verification) for any of your email accounts (e.g. official, personal or business email accounts etc.)?
- Yes
- No

(7) Have you ever used two-factor authentication (a.k.a. 2-FA or two-step verification) for any of your email accounts (e.g. official, personal, or business email accounts etc.)?
- Yes
- No

(8) For what kind of email account(s) do you use/have you used two-factor authentication (a.k.a. 2-FA or two-step verification)? Please select all that apply
- Personal account
- Office/business account
- School/College account
- Other (Please specify)
- I have never used two-factor authentication (a.k.a. 2-FA or two-step verification)

(9) Do you use a smartphone?
- Yes
- No

(10) What kind of smartphone do you use?
- iPhone
- Android
- Other (Please specify)
- I do not use a smartphone

(11) Do you use a personal computer (e.g., laptop/desktop etc.)?
- Yes
- No

(12) What kind of smartphone do you use?
- Windows
- Mac
- Linux
- Other (Please specify)
- I do not use a personal computer

A.1.2 Manipulation Check Questionnaire.

(1) Have you ever seen the particular image displayed in the story line before?
- Yes
- No

(2) Where did you see this image?

A.1.3 After Vignette Part 1.

(1) Bob did the right thing by not enabling two-step verification.
- Strongly disagree
- Somewhat disagree
- Neither agree nor disagree
- Somewhat agree
- Strongly agree

(2) Please explain the reasoning behind your answer.

A.1.4 After Vignette Part 2.

(1) Is it possible to launch an attack similar to the one presented in the story?
- Yes
- No
- I do not know

(2) How likely is it for someone to experience an attack similar to the one presented in the story?

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

(3) Enabling two-step verification (2-FA) would not have prevented the attack
- Strongly disagree
- Somewhat disagree
- Neither agree nor disagree
- Somewhat agree
- Strongly agree

(4) Please explain the reasoning behind your answer.
(5) Please rate how much you agree with each of the statements below. Note that all the statements below may not be applicable for Bob's scenario presented to you. Please select "Not Applicable" if a statement is not applicable for the presented scenario.
(Participants rated the items in the following scale: Not Applicable, Strongly disagree, Somewhat disagree, Neither agree nor disagree, Somewhat agree, Strongly agree)
- Bob should blame the attacker for the compromise.
- Bob should blame himself for not taking the responsibility for not enabling two-step verification (2-FA) for his email.
- Bob should blame himself for not taking the time to learn more about two-step verification (2-FA).
- Bob should blame himself for not being more careful about the data in his email account.
- Bob should blame the company for providing the service with security vulnerabilities.
- Bob should blame the email service provider for providing the service with security vulnerabilities.
- Bob should blame the company for failing to communicate risks adequately.
- Bob should blame the email service provider for failing to communicate risks adequately.
- The advertisement message promoting two-step verification (2-FA) should have been clearer and explained what the two-step verification (2-FA) was for.
- Bob should blame his friends for not informing him about two-step verifications (2-FA) and the attack.
- Bob should blame his colleagues for not informing him about two-step verifications (2-FA) and the attack.
- Bob should blame the government for not raising public awareness regarding the importance of using two-step verification (2-FA).

A.1.5 Demographic Questionnaire.

(1) What is your age (in years)?
(2) What is your gender?
- Male
- Female
- Other
- I prefer not to answer

(3) What is the highest level of education you have received?
- Less than high school
- High school graduate or GED
- Some college
- 2 year degree
- 4 year degree
- Master's degree
- Doctoral degree
- Professional degree

A.2 Storylines

A.2.1 Vignette Part 1.

Group 1A/1B: Official email - Intended use. Bob works as an insurance claims processor for a private health insurance company that has over 100,000 clients. Bob is provided an official email account by his company for official use (e.g., to exchange health insurance claim files). Bob always uses his office email account for official purposes only. On Dec 7, 2019, he received a message from his company (shown below) promoting a security feature called two-step verification (2-FA) that is supposed to enhance the security of the account. Bob decided not to activate the feature.

Group 2A/2B: Official email - Both official and personal use. Bob works as an insurance claims processor for a private health insurance company that has over 100,000 clients. Bob is provided an official email account by his company for official use (e.g., to exchange health insurance claim files). Bob always uses his official email account for official purposes, and sometimes for personal purposes. On Dec 7, 2019, he received a message from his company (shown below) promoting a security feature called two-step verification (2-FA) that is supposed to enhance the security of the account. Bob decided not to activate the feature.

Group 3A/3B: Personal email - Intended use. Bob works as an insurance claims processor for a private health insurance company that has over 100,000 clients. Bob is provided an official email account by his company for official use (e.g., to exchange health insurance claim files). Bob also has a separate personal email account, which he uses for personal purposes only. On Dec 7, 2019, he received a message from the personal email service provider (shown below) promoting a security feature called two-step verification (2-FA) that is supposed to enhance the security of the account. Bob decided not to activate the feature.

Group 4A/4B: Personal email - Both official and personal use. Bob works as an insurance claims processor for a private health insurance company that has over 100,000 clients. Bob is provided an official email account by his company for official use (e.g., to exchange health insurance claim files). Bob also has a separate personal email account, which he uses for personal purposes, and sometimes for official purposes. On Dec 7, 2019, he received a message from the personal email service provider (shown below) promoting a security feature called two-step verification (2-FA) that is supposed to enhance the security of the account. Bob decided not to activate the feature.

A.2.2 Vignette Part 2.

Group 1A/1B and Group 2A/2B: Official email. As you can recall from the first part of the story, on Dec 7, 2019, Bob decided not

Group	Number of Valid Participants	Age	Gender Breakdown
1A	33	Mean = 35.15, Median = 32, SD = 9.5	20 Male, 11 Female, 1 Other, 1 Prefer not to answer
1B	32	Mean = 40.63, Median = 41, SD = 12.54	19 Male, 12 Female, 1 Prefer not to answer
2A	32	Mean = 36.34, Median = 34, SD = 10.48	17 Male, 15 Female
2B	31	Mean = 42.3, Median = 41, SD = 12.05	17 Male, 14 Female
3A	32	Mean = 38.16, Median = 36.5, SD = 10.76	16 Male, 16 Female
3B	32	Mean = 39.78, Median = 37.5, SD = 9.99	18 Male, 14 Female
4A	31	Mean = 35.26, Median = 35, SD = 10.4	22 Male, 9 Female
4B	32	Mean = 43.75, Median = 44.5, SD = 11.23	11 Male, 21 Female
Total	255		

Table 3: Participant demographics by groups.

	Code	1A	1B	2A	2B	3A	3B	4A	4B	Total
Reason for agreement	Inconvenience	6%	3%	0%	6%	6%	22%	3%	9%	7%
	His choice/Personal decision	0%	3%	0%	6%	6%	25%	3%	6%	6%
	Scam	0%	0%	0%	0%	3%	0%	13%	0%	2%
Reason for disagreement	Benefits of 2-FA	63%	47%	56%	19%	72%	31%	35%	41%	46%
	Responsibility towards company/Client data	24%	47%	41%	52%	3%	0%	19%	31%	27%
	Peace of mind	9%	3%	3%	0%	3%	0%	0%	0%	2%
	Vague/Unable to code	9%	9%	16%	6%	6%	3%	10%	9%	9%
	Total valid responses	33	32	32	31	32	32	31	32	255

Table 4: Codes for the statement: "Bob did the right thing by not enabling two-step verification (2-FA)".

	Code	1A	1B	2A	2B	3A	3B	4A	4B	Total
Reason for agreement	Server attack would not have been prevented	18%	9%	13%	16%	16%	7%	6%	3%	11%
	Skeptical about effectiveness of 2-FA	33%	25%	38%	35%	19%	31%	16%	31%	29%
Reason for disagreement	2-FA would have been effective for email	48%	56%	56%	48%	69%	53%	71%	47%	56%
	Early notification/Warning on cell phone	6%	3%	16%	3%	9%	3%	10%	0%	6%
	Additional step/Require access to Bob's cell phone	15%	28%	28%	16%	38%	16%	29%	25%	24%
	Requires everyone to use 2-FA	0%	0%	6%	0%	0%	0%	0%	0%	1%
	Vague/Unable to code	15%	16%	6%	23%	9%	13%	16%	22%	15%
	Total valid responses	33	32	32	31	32	32	31	32	255

Table 5: Codes for the statement: "Enabling two-step verification (2-FA) would not have prevented the attack".

to activate two-step verification. On Dec 15, 2019, security attackers broke into the company email authentication server and stole login credentials of several thousand employees, including Bob's credentials. However, Bob was unaware of the attack. On Dec 16, 2019, the attacker used the stolen credentials to log in into Bob's official email account and change the password, preventing him from accessing his own email account.

Group 3A/1B and Group 4A/4B: Personal email. As you can recall from the first part of the story, on Dec 7, 2019, Bob decided not to activate two-step verification. On Dec 15, 2019, security attackers broke into the personal email service provider's email authentication server and stole login credentials of several thousand users, including Bob's credentials. However, Bob was unaware of the attack. On Dec 16, 2019, the attacker used the stolen credentials to log in into Bob's personal email account and change the password, preventing him from accessing his own email account.

A.3 Demographics Summary

Table 3 shows group-wise summary of participants' demographics.

A.4 Qualitative Codes across the Groups

Qualitative codes for the statements, "Bob did the right thing by not enabling two-step verification (2-FA)" and "Enabling two-step verification (2-FA) would not have prevented the attack" are presented in Table 4 and 5 respectively.

Finding	Implication
Adopters indicated more blame towards the user (i.e., Bob) compared to non-adopters who were more inclined to hold external entities (e.g., service provider, the promotional message) more responsible.	1. Based on the relationship between blaming tendency and coping response behavior shown in social cognitive frameworks [17, 66], adopters, with their "self-blame" tendency, are more likely to exhibit positive coping response and follow the recommended actions after a breach than the non-adopters who are likely to show poor-coping and hold others responsible for the attack. 2. A data breach incident is likely to cause feelings of breach of trust substantially more among the non-adopting group compared to adopting group, with the belief that service providers are responsible for user security; requiring extensive post-breach calibration of cognitive and affective processing as part of post-breach communication.
Non-adopters blamed the service provider more for the attack, indicating a belief that service providers are more responsible for keeping users secure.	1. Pre-breach messaging should underscore the importance of shared responsibility that can change the perception of self-responsibility and the current dynamics between the end-users and service providers (e.g., being mindful of software update behavior or phishing attack following a security training).
Non-adopting population noted the technical feasibility of the attack to be significantly less than adopters, indicating a low perceived vulnerability of cyberattacks.	1. Information campaigns should target users' low perceived vulnerability and incorporate such messaging while promoting cybersecurity tools/behavior, both in the context of pre and post-breach communication.
Attribution of blame was shifted significantly based on the account types and usage scenarios explored across the groups.	1. Underscores the importance of considering contexts while designing promotional messages for communicating risks in pre and post-breach scenarios. 2. Service provides should put importance on context-sensitive promotional messages for promoting security tools/behavior that consider users' specific usage patterns.
Both technical and non-technical participants showed skepticism about the efficacy of 2-FA in a similar way.	1. Skepticism regarding the efficacy of security tools/features (e.g., 2-FA) can lead to a sense of "hopelessness", and plays a role in users' non-compliant behavior. Hence, pre-breach communications of a security tool should address any such skepticism, while explaining the responsibility of the users' as well. 2. As users are likely to feel betrayed by a security breach, post-breach communication should explain why the existing security features failed to prevent the attack and how the recommended action will minimize the risk in the future.

Table 6: Summary of key implications for pre-breach and post-breach communications based on our findings.

A.5 Summary of Implications for Pre and Post-breach Risk Communication

Table 6 presents a summary of the implications derived from the study results for pre-breach and post-breach security risk communications.

A.6 2-FA Promotional Message

Figure 2 and 3 show the original 2-FA promotional message by Google and the edited version that was used in the study, respectively.

A.7 Frequency Distribution Graphs - Blaming Bob

Figure 4, 5 and 6 shows distribution graphs for participants' blame attribution towards Bob. The findings are presented in Section 4.1.2.

A.8 Frequency Distribution Graphs - Blaming the Entity

Figure 7, 8, 9, and 10 shows distribution graphs for participants' blame attribution towards the entity. The findings are presented in Section 4.1.3.

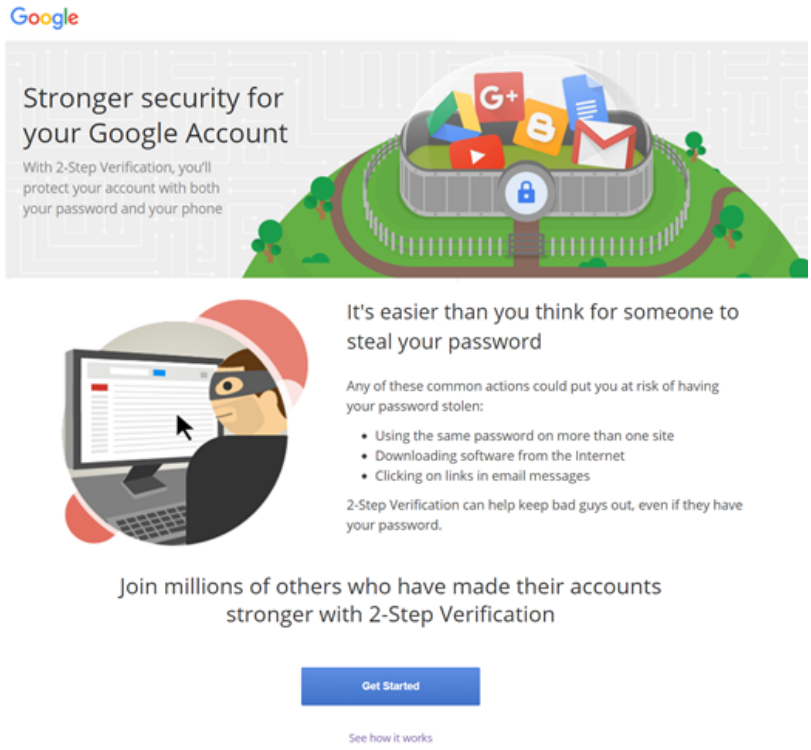


Figure 2: Google's original 2-FA image.

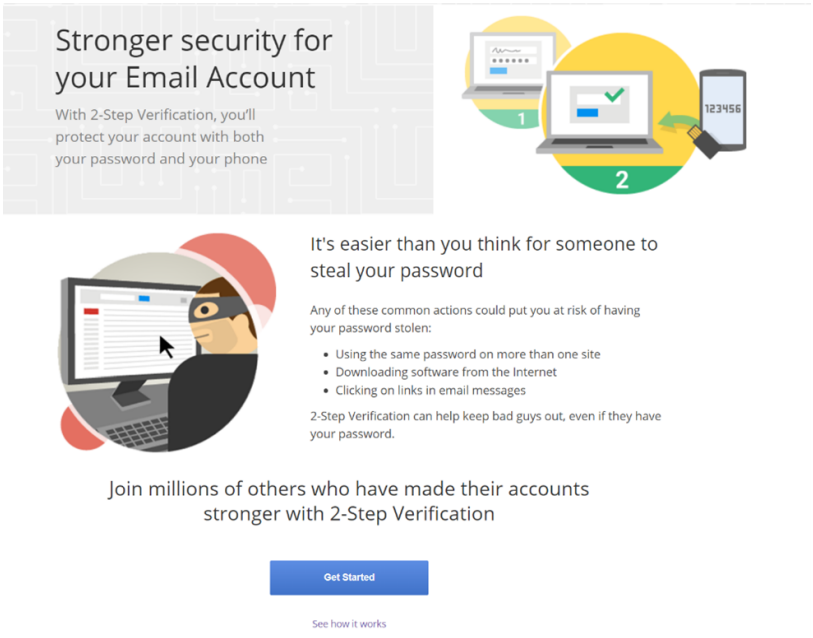
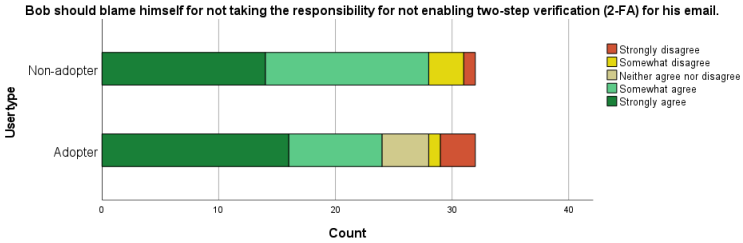
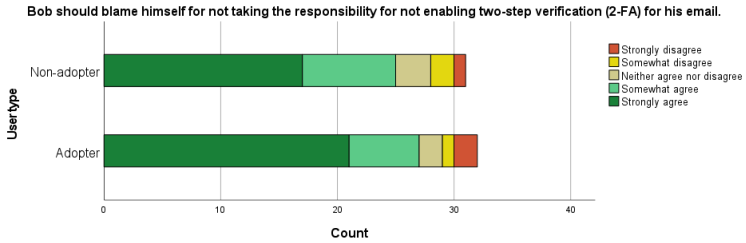


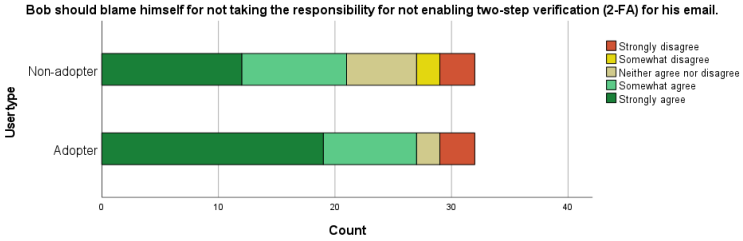
Figure 3: Edited 2-FA image used in the study.



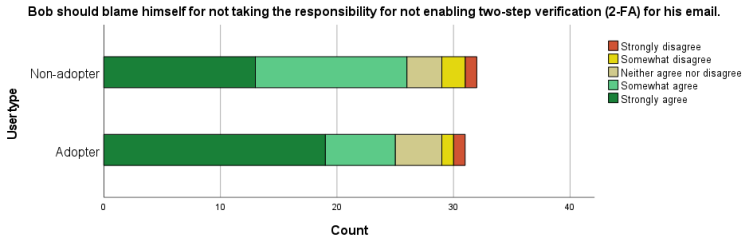
(a) Official email for official purpose scenario



(b) Official email for both purposes scenario

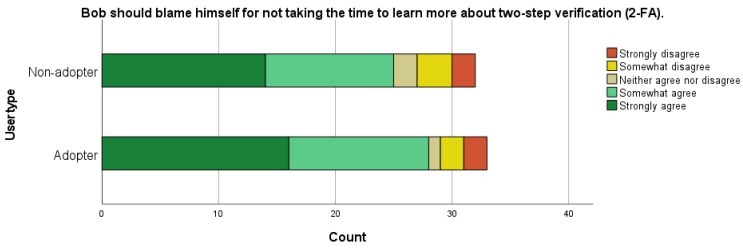


(c) Personal email for personal purpose scenario

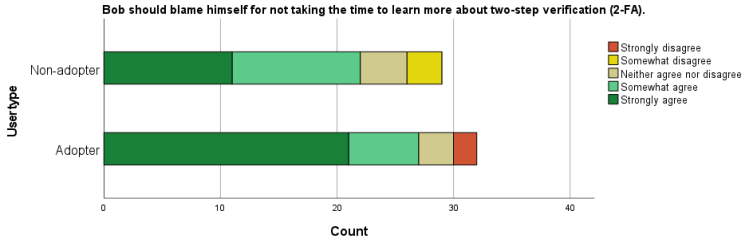


(d) Personal email for both purposes scenario

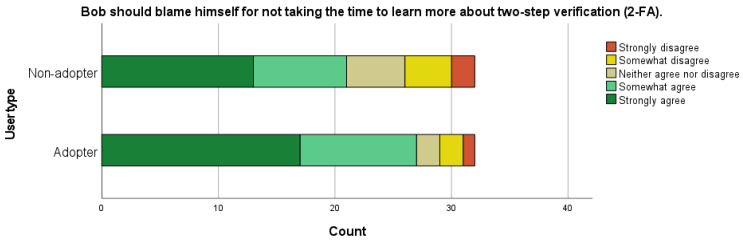
Figure 4: Response distributions representing the item "Bob should blame himself for not taking the responsibility for not enabling two-step verification (2-FA) for his email" among adopters and non-adopters of 2-FA.



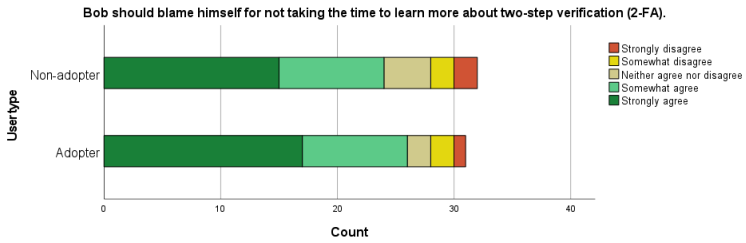
(a) Official email for official purpose scenario



(b) Official email for both purposes scenario

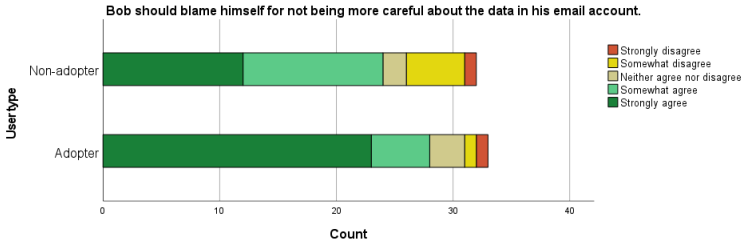


(c) Personal email for personal purpose scenario

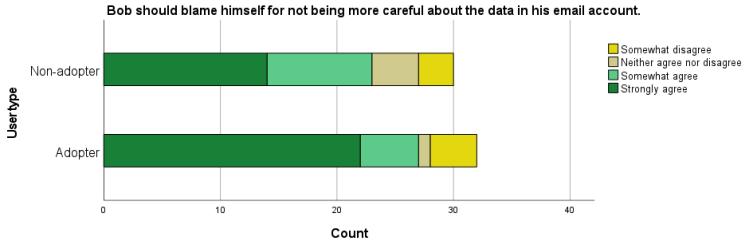


(d) Personal email for both purposes scenario

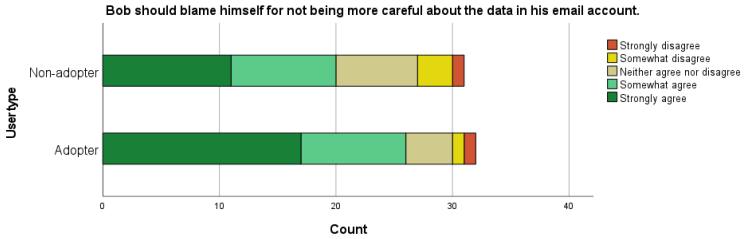
Figure 5: Response distributions representing the item "Bob should blame himself for not taking the time to learn more about two-step verification (2-FA)" among adopters and non-adopters of 2-FA.



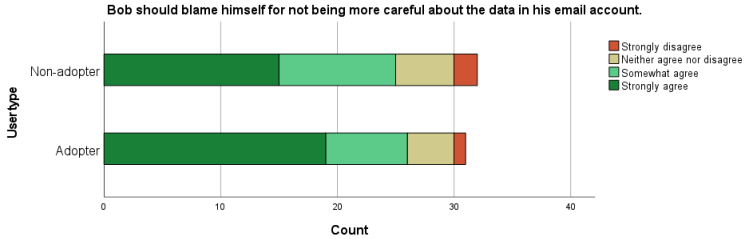
(a) Official email for official purpose scenario



(b) Official email for both purposes scenario

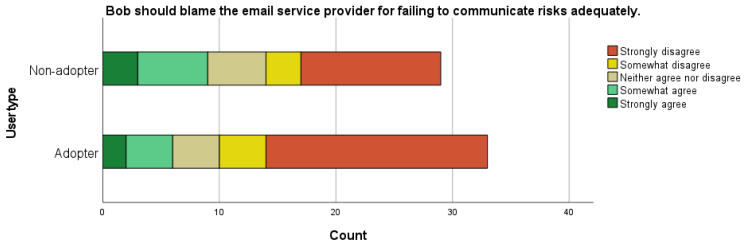


(c) Personal email for personal purpose scenario

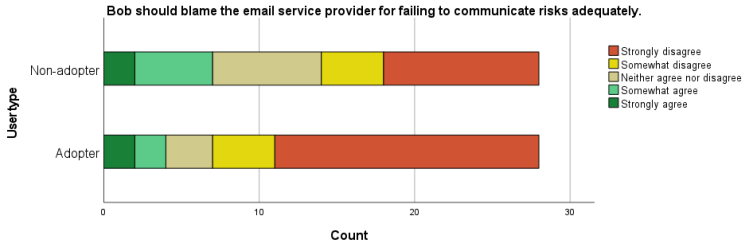


(d) Personal email for both purposes scenario

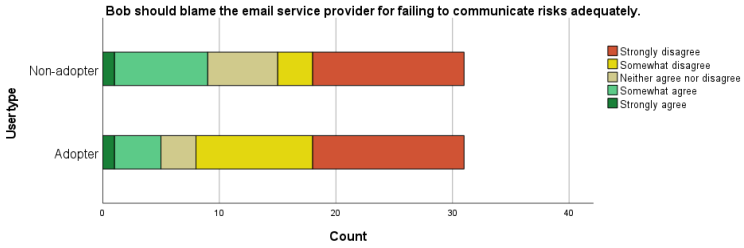
Figure 6: Response distributions representing the item "Bob should blame himself for not being more careful about the data in his email account" among adopters and non-adopters of 2-FA.



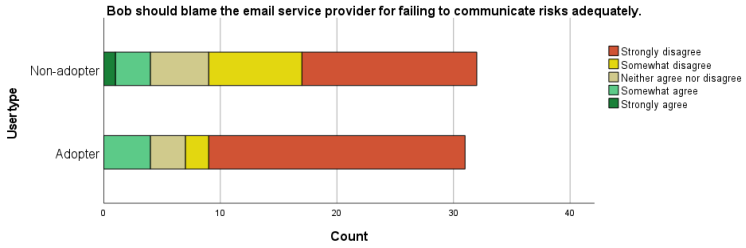
(a) Official email for official purpose scenario



(b) Official email for both purposes scenario

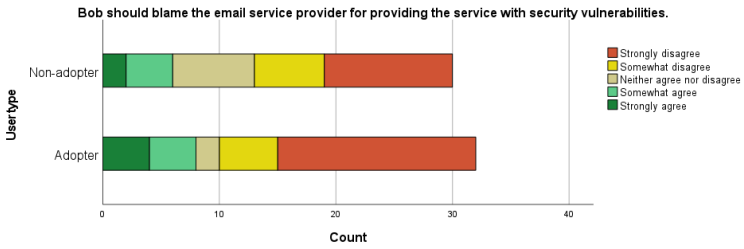


(c) Personal email for personal purpose scenario

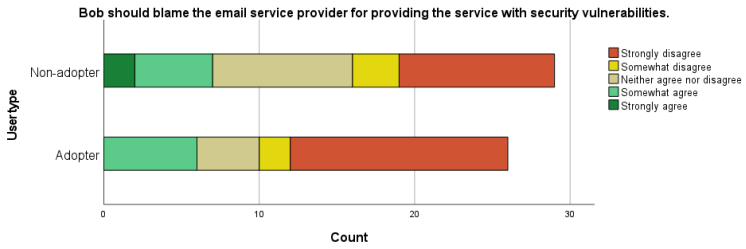


(d) Personal email for both purposes scenario

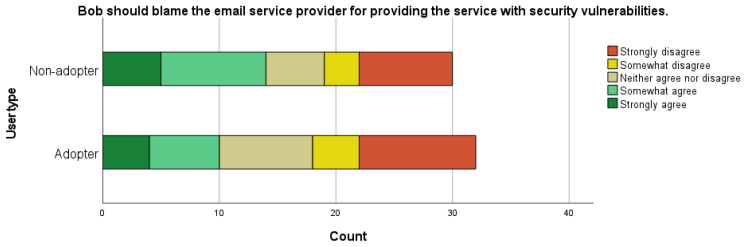
Figure 7: Response distributions representing the item "Bob should blame the email service provider for failing to communicate risks adequately" among adopters and non-adopters of 2-FA.



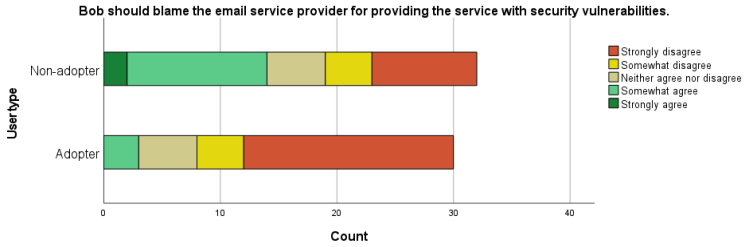
(a) Official email for official purpose scenario



(b) Official email for both purposes scenario

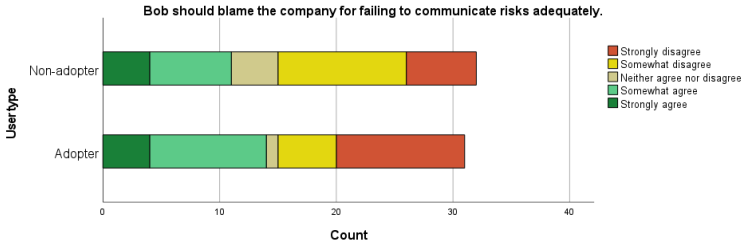


(c) Personal email for personal purpose scenario

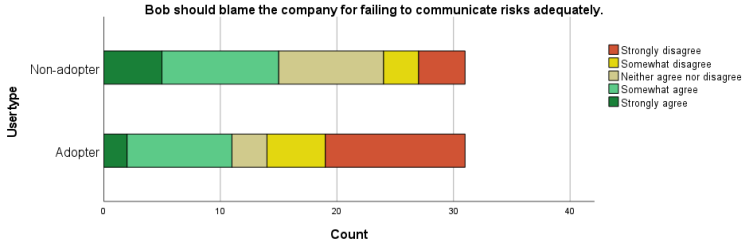


(d) Personal email for both purposes scenario

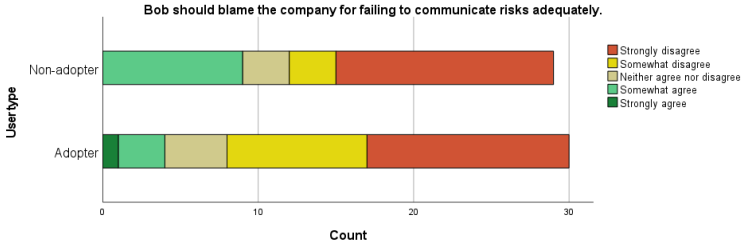
Figure 8: Response distributions representing the item "Bob should blame the email service provider for providing the service with security vulnerabilities" among adopters and non-adopters of 2-FA.



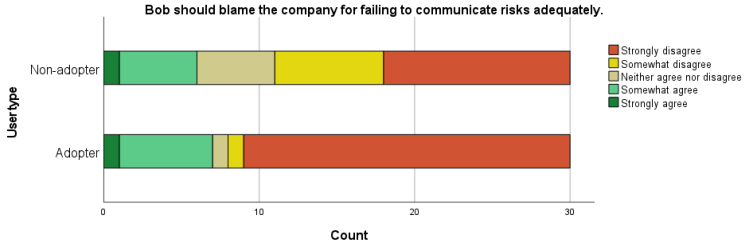
(a) Official email for official purpose scenario



(b) Official email for both purposes scenario

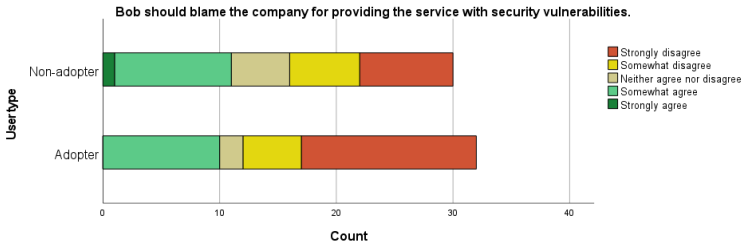


(c) Personal email for personal purpose scenario

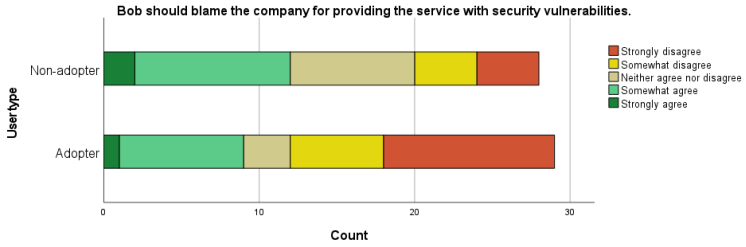


(d) Personal email for both purposes scenario

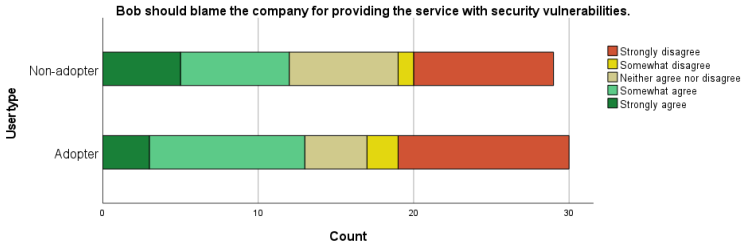
Figure 9: Response distributions representing the item "Bob should blame the company for failing to communicate risks adequately" among adopters and non-adopters of 2-FA.



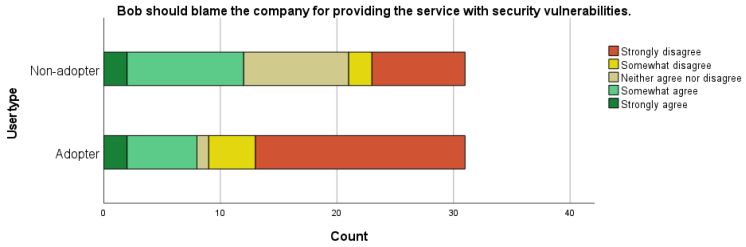
(a) Official email for official purpose scenario



(b) Official email for both purposes scenario



(c) Personal email for personal purpose scenario



(d) Personal email for both purposes scenario

Figure 10: Response distributions representing the item "Bob should blame the company for providing the service with security vulnerabilities" among adopters and non-adopters of 2-FA.