Understanding the Viability of Gmail's Origin Indicator for Identifying the Sender

Enze Liu, *UC San Diego*
Lu Sun, *UC San Diego*
Alex Bellon, *UC San Diego*
Grant Ho, *University of Chicago*
Geoffrey M. Voelker, *UC San Diego*
Stefan Savage, *UC San Diego*
Imani N. S. Munyaka, *UC San Diego*

Abstract

The current design of email authentication mechanisms has made it challenging for email providers to establish the authenticity of email messages with complicated provenance, such as in the case of forwarding or third-party sending services, where the purported sender of an email is different from the actual originator. Email service providers such as Gmail have tried to address this issue by deploying sender identity indicators (SIIs), which seek to raise users' awareness about where a message originated and encourage safe behavior from users. However, the success of such indicators depends heavily on user interpretation and behavior, and there exists no work that empirically investigates these aspects. In this work, we conducted an interactive survey (n=180) that examined user comprehension of and behavior changes prompted by Gmail's passive SII, the *'via'* indicator. Our quantitative analysis shows that although most participants (89%) noticed the indicator, it did not have a significant impact on whether users would adopt safe behaviors. Additionally, our qualitative analysis suggests that once prompted to consider why *'via'* is presented, the domain name displayed after 'via' heavily influenced participants' interpretation of the message *'via'* is communicating. Our work highlights the limitations of using passive indicators to assist users in making decisions about email messages with complicated provenance.

1 Introduction

Email is perhaps the longest-lived service in continuous use on the Internet and its precursor networks — dating back to at least Tomlinson's SNDMSG in 1971. As a result, email standards have not enjoyed the luxury of a careful design, but have instead accreted new mechanisms to shore up the legacy Simple Message Transfer Protocol (SMTP) against newfound problems. Chief among these problems has been email spoofing, whereby an adversary sends messages purporting to be from an address that does not, in fact, belong to them (e.g., for spam, phishing, etc.). To help mitigate such abuse, email protocol designers have added a range of out-of-band authentication protocols — SPF, DKIM and DMARC, among others — to help validate the identity of the sending organization (i.e., domain name) in an email message.

However, these mechanisms are hindered in practice because of modern Internet email borrowing heavily from the practices of mid-20th century business correspondence, including the notions of "carbon copies" (cc), message forwarding, and distribution lists.[1] In particular, both email forwarding and distribution lists require that messages be distributed by a third-party who is not the original sender — highly similar to spoofing. Thus, there are a range of legitimate scenarios where existing email authentication protocols will fail to validate the identity of the sender. To deal with this ambiguity, many email service providers (e.g., Google's Gmail and Microsoft's Outlook 365) choose to prioritize *deliverability* over possible security threats and will allow many such messages to reach user mailboxes [51].

This situation leaves individual users with the burden of distinguishing spoofed email messages from those messages that were merely ambiguously sourced. Moreover, the standard information displayed by a Mail User Agent (MUA) (e.g., To:, From:, Subject:, Date:, etc.) does not provide any indication that such a situation is even present, let alone provide sufficient evidence for making an informed decision. Spero and Biddle identify this issue as well, opining that "making the Mail-from (the true origin of the email) more visible would be beneficial, along with some information about the Mail-from domain" [74]. Gmail is one of the only two MUAs that attempts to inform their users of such situations, by providing a 'via' indicator in its user interface. Thus, a message from "alice@foo.com via bar.com" is intended to convey that the message claims to originate from `foo.com`, but was actually delivered by `bar.com`.

However, the utility of this indicator depends on the extent to which users understand its meaning, intuit its purpose, and are able to apply that understanding to then make informed

[1] The incorporation of these norms into email systems dates back at least to 1978, with Shoen's 1978 Mail client distributed with BSD Unix.

choices. In this paper, we examine the utility of Gmail's 'via' indicator to answer the following questions:

RQ1 How do users respond to the 'via' indicator?

RQ2 How would users react to the email when the 'via' indicator is present?

RQ3 What message is the indicator communicating to end users?

RQ4 What are users' perceptions of the relationship between the two domains shown by the indicator?

We answer these questions by surveying Prolific gig workers about the Gmail indicator. We replicated the Gmail interface, and asked participants to interact with a message from "alerts@chase.com" and answer follow-up questions about their experience. We employ a mixed-methods approach to our analysis to understand their interpretations of the 'via' indicator and identify how users respond to Gmail providing the indicator. We consider our results to be an upper bound baseline since gig workers are often more skilled in using various technologies.

Our results suggest that even with years of email experience, the 'via' indicator is not a factor in users' email decision-making process. Most of the participants (89%, n=120) that were shown the indicator remembered seeing it during the study. However, even in the case where the domain name displayed after 'via' (hence referred to as the 'via' domain) was `rlxaz.xyz`, most participants (85%, n=60) still believed Chase Bank or `chase.com` was the sender. Among these participants, 78% of them were "very confident" about their answers. Our results also suggest that the 'via' domain directly impacts users' interpretations of the indicator's purpose. In particular, users believed that the email they viewed was coming from `chase.com` through another part of the Chase Bank business when the 'via' domain was `chasesupport.com`.

These findings suggest that passive indicators that rely on user interpretation are likely to have limited success. In our study, once participants were asked to meditate on the purpose of 'via' from different perspectives, their interpretation evolved such that some participants completely changed their interpretation of the email. We suggest that future sender identity indicators be designed to communicate the necessary information users need without additional prompting. Ultimately, we make the following contributions:

- We provide an overview of how end users interpret the 'via' indicator and the factors that influence these beliefs.
- We present one of the first comparisons of user behavior in response to the indicator, a result that complements prior research on warning design and phishing susceptibility research.
- We identify challenges in communicating sender identity to technically experienced users and discuss how these barriers increase user risk.

2 Related Work

Our work falls under the domain of phishing prevention and email spoofing. We start by reviewing the prior literature on phishing prevention and then discuss relevant literature on email spoofing.

2.1 Phishing Prevention

Prior work on phishing prevention focuses on three main areas: (1) understanding users' phishing susceptibility and improving phishing training; (2) automatically detecting phishing attacks without user interaction; and (3) warning users about potential risks.

2.1.1 Phishing Susceptibility and Training

Because phishing exploits human mistakes rather than software vulnerabilities [98], researchers have investigated the reasons why users fall for phishing attacks and how to improve anti-phishing training and educational material. Prior work has found a variety of tactics that can make phishing email messages more persuasive [6, 9, 14, 30, 50, 57, 60, 79, 91], including having recognizable logos, targeting recipients' specific contexts, and using persuasive techniques, among others. Similarly, papers have identified a wide range of factors that affect users' susceptibility to phishing attacks [8, 9, 16–19, 29, 36, 50, 57, 59, 60, 71, 78, 79, 84, 86, 92], such as their age, personal traits, prior training, gender, and strategies they employ to detect phishing email messages. Leveraging these insights, other studies have focused on improving anti-phishing training [18, 19, 36]. This prior work includes exploring the efficacy of different training formats such as embedded training [11, 42, 45], teaching anti-phishing via games [45, 72, 87], and how the effectiveness of training varies in different contexts [37, 41, 43, 44, 46, 64, 66, 73, 85].

2.1.2 Automated Detection

Automated detection systems serve as the first line of defense by identifying attacks before users see them. The community has used a variety of algorithms to detect phishing email messages, websites, and URLs. These approaches range from commercial spam filters [63] to heuristics [13, 34, 40] and machine learning models [1, 20, 25, 75] proposed in academic work.

To detect attacks, these algorithms extract features from an email message, URL, and/or website and then apply a set of rules or machine learning model to identify phishing attacks. Prior work has explored a variety of different feature

sets [27, 54, 55, 83, 89, 94] and algorithms [38, 88, 99] to improve detection accuracy. Finally, simple approaches such as blocklists of IP addresses and accounts [7, 28, 53, 62] are also widely deployed in practice for phishing detection.

While beneficial, these detection systems face practical limitations. They can produce a large number of false positives when deployed at scale due to the high volume of benign email messages [61]. They also can be evaded by sophisticated adversaries [31]. As a result, automated phishing detection algorithms are often paired with phishing warnings to improve their effectiveness [45, 61].

2.1.3 Phishing Warnings and Indicators

Phishing warnings and indicators complement automated detection systems by alerting users of potential risks and supplying additional information to help users make informed decisions. Prior work has proposed different kinds of phishing warnings, including Passpet [96], dynamic security skins [15], SpoofGuard [77], Trustbar [33], social saliency nudges [58], active warning dialogs [10], and phishing warnings employed by browsers such as Chrome and Firefox [2]. Past research on these indicators has shown that passive indicators such as security toolbars are ineffective [22, 93], and active indicators that interrupt a user's current task are more useful in practice [22]. There also exists ongoing research that investigates the effectiveness of different anti-phishing support systems [68] as well as how to better design inclusive email security indicators [97].

Beyond these high-level warnings, other work has found that even subtle warning design choices can have a noticeable impact on the efficacy of phishing warnings and indicators [26]. In terms of ineffective warning design, prior work has found that user habituation to warnings [22] and failure to present information in a succinct and understandable fashion [16,93] lead to poor warning efficacy. While Lin et al. [49] report that using only domain highlighting as a browser warning does not provide strong protection against phishing, Volkamer et al. [80] found that combining domain highlighting with forced attention to a browser's address bar largely improves phishing detection. They also noted, in a separate study [81], the potential benefits of providing just-in-time and just-in-place tooltips, which follow-up studies [61] have confirmed. Zheng et al. [100] investigated the (in)effectiveness of presenting users with full email header details. Examining the use of multiple defenses, Yang et al. [95] discovered through a field experiment that combining phishing training and active phishing warnings can significantly reduce the click-through rate.

2.2 Email Spoofing

The original design of the Simple Mail Transfer Protocol (SMTP) lacked authentication, making email spoofing both possible and common [51]. Prior work examines a range of techniques attackers can use to successfully send spoofed email messages. Hu et al.'s [35] measurement study showed that many major mail providers delivered spoofed email to user inboxes without noticeable errors or warnings, and Chen et al. [12] demonstrated how attackers can compose multiple inconsistencies in different mail servers and clients to reliably send a spoofed email. More recently, several papers have explored how attackers can abuse email forwarding to send spoofed email messages. This work includes Shen et al.'s [70] large scale analysis on email spoofing attacks, Wang et al.'s [82] study on email spoofing opportunities introduced by Authenticated Receiver Chain (a standard for verifying servers that forward email), and Liu et al.'s [51] study on attacks enabled by email forwarding. However, all efforts mentioned above focused on the technical aspects of email spoofing. Our work is one of the first to examine the effectiveness of Gmail's 'via' indicator designed to mitigate such attacks.

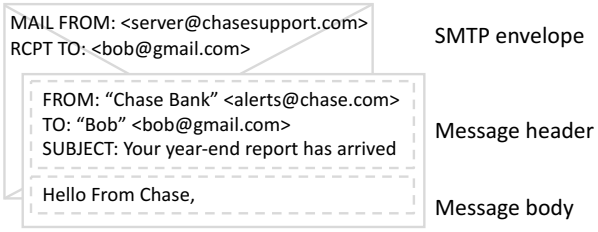


Figure 1: An example of SMTP headers, inspired by Figure 3 from Chen et al. [12].

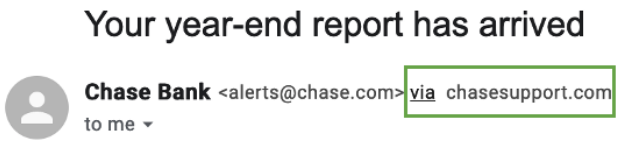
3 Background

In this section, we give a brief overview of SMTP (the protocol which governs the transmission of email) and provide background on sender identity indicators.

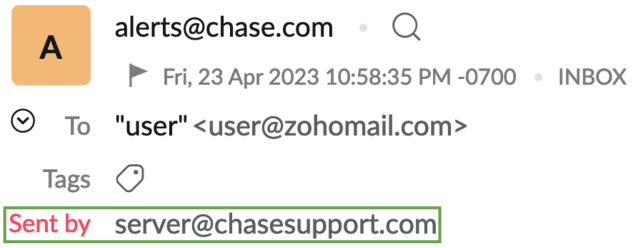
3.1 Simple Mail Transfer Protocol

Under the Simple Mail Transfer Protocol (SMTP), an email message includes two sets of headers that represent the sender(s) and recipient(s) of an email. Figure 1 shows an example message with both sets of headers. One set of headers, the SMTP envelope headers, consists of the MAIL FROM field and the RCPT TO field, and provides email servers with routing and delivery instructions. Specifically, the MAIL FROM field specifies the server that sent the email (`server@chasesupport.com`), and the RCPT TO field specifies the recipient of the email (`bob@gmail.com`).

The other set of headers, the SMTP message headers, includes the FROM and TO headers. This set of headers is used for user interface purposes only and does not affect email



(a) Gmail's SII



(b) Zoho's SII

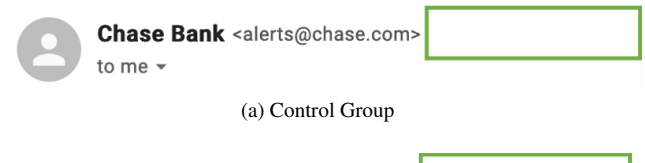
Figure 2: SIIs deployed by Gmail and Zoho with the SII highlighted.

routing [51]. Both the FROM and TO headers consist of a human-readable name and email address. In Figure 1, the FROM header consists of the human-readable name "Chase Bank" and the email address `alerts@chase.com`, and the TO header consists of the name "Bob" and email address `bob@gmail.com`.

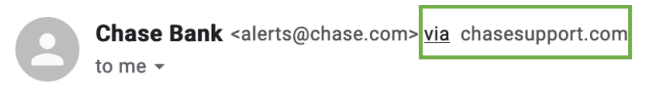
3.2 Sender Identity Indicators

Under the SMTP protocol, the MAIL FROM and RCPT TO headers are opaque to users, and users only see the information in an email's FROM and TO headers. This design works well when the MAIL FROM and FROM headers share the same domain. In practice, however, the domains in an email's MAIL FROM and FROM headers do not always match. This mismatch occurs for a range of both benign and malicious reasons, including email forwarding (e.g., by mailing lists) and third-party sending email services, as well as email spoofing. To address the issue of header spoofing, the community has developed defensive protocols such as SPF and DMARC [21], where domains can provide information to recipients that allow them to validate if an email message truly originated from the domain, and that specify actions to take if such validation fails.

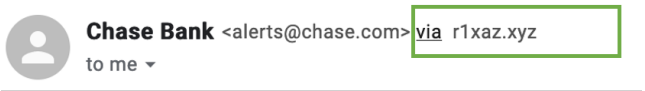
Unfortunately, due to limitations in these protocols and the lack of universal adoption, many recipient email servers cannot robustly authenticate all email messages. Moreover, in an effort to prioritize email deliverability [51], many domains often specify a permissive policy for recipients to follow if an email message fails to authenticate under a protocol like SPF or DMARC. As a result, major email providers such as Gmail and Microsoft Outlook often deliver email messages of unknown or potentially questionable authenticity.



(a) Control Group



(b) Support Group



(c) Random Group

Figure 3: Headers of the email shown to participants in each group.

To mitigate some of these issues, two email providers (Gmail and Zoho) have introduced UI modifications designed to provide additional information and awareness to users about an email's potential origins. We refer to these UI features as "sender identity indicators" (SIIs). Figure 2 shows an example of their SIIs with the indicator highlighted. In our work, we focus exclusively on Gmail's SII, the 'via' indicator, given Gmail's wide adoption [52]. Gmail uses 'via' to display the actual originator of an email message to recipients. For this message, the purported sender is `chase.com`, yet the actual originator is `chasesupport.com`. Once again, this mismatch can be due to using third-party sending services in a benign case, or email spoofing in a malicious case. Gmail cannot distinguish between the two cases, delegating the risk and leaving the decision up to the recipient (with the indicator as an aid).

4 Methodology

We use a between-subject study design to observe how the presence of the 'via' indicator impacts users' perception of the email sender when viewing an email in the Gmail interface. To answer these questions, we design a replica Gmail interface and survey participant groups under three different email conditions: "Control", "Support" and "Random". Participants in the *Control* group are presented with an email that has no 'via' indicator (Figure 3a). Participants in the *Support* group are presented with an email with the 'via' indicator (Figure 3b), which is followed by the `chasesupport.com` domain. Participants in the *Random* group are presented with an email with the 'via' indicator, which is followed by the `r1xaz.xyz` domain (Figure 3c). The Support group simulates a situation where the 'via' domain resembles the target domain, while the Random group simulates a situation where the 'via' domain is an unfamiliar domain. Participants were asked to log into a web interface modeled after Gmail, locate

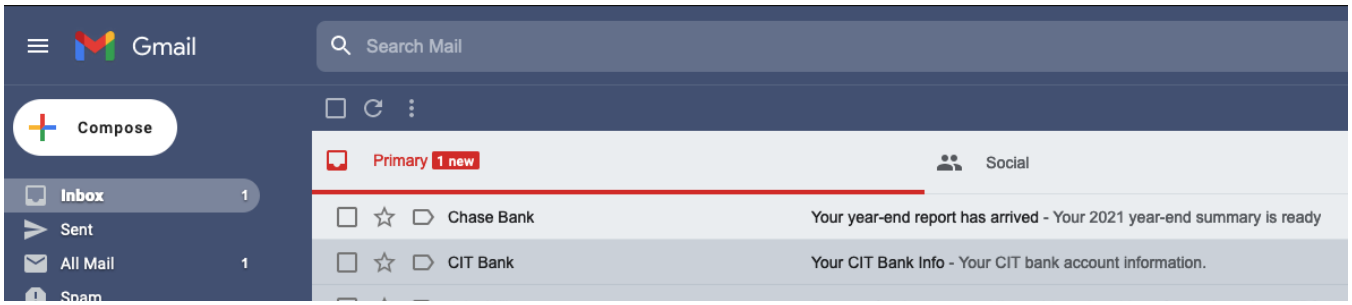


Figure 4: Email client landing page

and examine a specific email message, and answer questions about the email (Section 4.2). We randomly allocated participants to the three groups, which we will refer to below as Support, Random, and Control. We limit our scenarios to one email message and two 'via' domains to reduce the number of variables and focus specifically on understanding how people respond to 'via'. Below, we start by providing a brief description of our web-based email client and the email shown to the participants (Section 4.1), followed by a detailed description of our survey design and analysis methods.

4.1 Email Client

We built a web-based email client that is modeled after Gmail's web client. Our study focuses on Gmail because it is the most widely-used mail provider [52]. We decided to build a replica of the Gmail client instead of sending spoofed email messages to participants' real accounts so that we could easily track participants' interaction with the email in a controlled environment and avoid crossing ethical research boundaries.

Figure 4 shows the landing page of our web client. This page presents a list of email messages to the user, and we highlight the email that they need to review. We did not remove the Gmail brand name, as we seek to simulate users' experience with Gmail's web interface and increase ecological validity.

Upon clicking on the email that they are asked to review, participants are shown a page that displays the content of an email. This page mainly consists of two parts: email headers and email content. Figure 3 shows the email headers displayed to each of the survey groups. Users also have access to detailed header information that would be available in Gmail's web interface by clicking on the gray down-arrow button, also shown in Figure 3. Figure 5 shows the actual email content displayed. We take the content from a real email message sent by Chase that contains a link (the view my summary button) but change the link address to the main Google search page to prevent negatively impacting participants.

Lastly, we track if any of the buttons are clicked, if the link in the email is clicked, and when and how long users browsed the web interface.

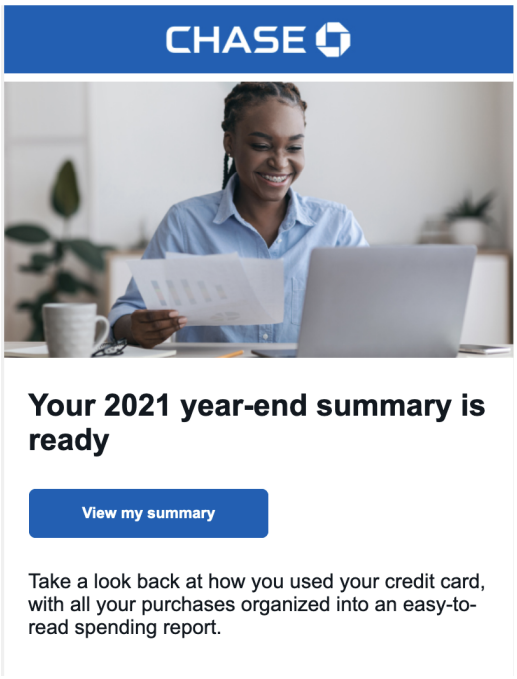


Figure 5: The email that participants needed to review.

4.2 Survey Protocol

We used Prolific to conduct our surveys. To avoid priming users for security, we framed the research as a study on the usability of the Gmail interface, including whether users are able to find an email and identify the sender of that email.

Since our study focuses on Gmail, we used a prescreening process to only include users with Gmail accounts. Specifically, we highlight in our survey description that participants must be Gmail users to enter the study. At the beginning of our survey, we also ask participants to confirm that they are indeed Gmail users and provide an option to exit the survey if they are not.

We give each user a unique link to our web-based email client (Section 4.1) after they pass prescreening. We embedded the link in the Qualtrics survey and instructed users to click on the link to access the email client in a separate tab,

Table 1: Demographics of survey participants

	Support *N* (%)	**Random** *N* (%)	**Control** *N* (%)
Age			
18-30	25 (42%)	28 (46%)	21 (35%)
31-50	30 (50%)	23 (39%)	30 (50%)
51-65	5 (8%)	7 (12%)	9 (15%)
Over 65	0 (0%)	2 (3%)	0 (0%)
Gender			
Female	22 (37%)	32 (53%)	28 (47%)
Male	36 (60%)	28 (47%)	32 (53%)
Non-binary	2 (3%)	0 (0%)	0 (0%)
Education			
No College degree	18 (31%)	22 (36%)	24 (40%)
2 year degree	5 (8%)	3 (5%)	2 (3%)
4 year degree	29 (48%)	24 (40%)	23 (38%)
Postgrad/Prof	8 (13%)	10 (17%)	11 (18%)
Prefer not to select	0 (0%)	1 (2%)	0 (0%)

and then return to Qualtrics to continue the study. In the survey, we started by asking users to imagine the email client was the actual Gmail web interface. Next, we instructed them to find, open, and read the email that was titled "Your year-end report has arrived". After this, we asked a series questions hosted with Qualtrics about the email (more details below). Participants had access to the email client throughout the study.[2]

First, we asked users to indicate the actions they would like to perform with the presented email by selecting from a list of available choices. This list of choices is adopted from prior work [18] and includes:

- Keep, save, or archive the email
- Click on the "View my summary" button in the email
- Forward the email to someone else
- Reply by email
- Contact the bank in other ways than email
- Delete the email
- Search a term in Google (please specify)
- Other (please specify)

We consider users who suggested that they would click on the link as having the potential to fall for phishing attacks, regardless of other actions they indicated.[3]

[2]Our survey questions, together with our implementation of the email client, can be found at https://github.com/ucsdsysnet/soups23-email-origin-indicator.

[3]While we did not have a follow-up phishing page that asked users to enter sensitive information, prior literature [41, 42, 71] has consistently suggested that 90% of the users who would click on the link would provide information on the phishing page.

Next, we asked users to answer three questions about the sender of the email: (1) the name of the person or entity that sent the email; (2) the email address of the person or entity that sent the email; and (3) how they decided the answer to the previous two questions. We also asked them to indicate their confidence level for questions (1) and (2) on a scale of 1 (not confident) to 5 (very confident).

We then moved on to ask users questions about the 'via' indicator. Specifically, we asked them to recall whether they saw the 'via' indicator during the study and whether they had encountered the 'via' indicator in the past before the study. We also asked them to indicate whether they understood what 'via' meant. For users who indicated that they knew the meaning of 'via', we followed up with a question asking them to explain what 'via' meant. For others who indicated that they did not know the meaning of 'via', we asked them to guess what information 'via' was trying to communicate. Lastly, for all users, we asked them to reflect on why Gmail chose to display the 'via' indicator.

Our last question probes the judgment made by users after having their attention directed to the 'via' indicator. We asked users to indicate whether they agreed that Chase.com used or instructed the 'via' domain (`rlxaz.xyz` or `chasesupport.com`) to send the email, and elaborate on their answer.

After answering the above questions and a demographic survey, users were debriefed about the true intention of this study and provided an option to have their data removed. We then thanked them for their participation and compensated them with $2.50 ($15/hr USD) for the 10 minute survey. We acquired approval from our institution's review board (IRB) before conducting the study.

4.3 Participants

Our sample size was informed by an a priori power analysis conducted with G Power [23] to determine the sample size needed for an effect size of .25 and alpha of .05 to test if one mean is significantly different among three groups. The results suggested a sample of 159 participants with 53 participants in each group for a power of .8. We received 180 unique responses across our three surveys, with 60 participants in each survey group. Most of our participants were between 31 and 50 years of age (46%), Male (53%), White (81%), and had a 4-year degree (42%). We compare the demographics of our participants, shown in Table 1, to the most recent US and UK Census data to evaluate how well they represent the US and UK populations. We saw that participants skewed toward younger age ranges than the US and UK populations, and higher educational attainment than the US population (but about the same as the UK population), meaning they were likely more familiar with computing concepts and usage.

In Table 2, we describe how familiar participants were with computers and phishing. The vast majority of participants

Table 2: Computer and email expertise demographics of survey participants

	Support Group *N* (%)	**Random Group** *N* (%)	**Control Group** *N* (%)
Computer Familiarity			
Work in or hold a degree in CS/IT	12 (20%)	6 (10%)	8 (13%)
Do not work in or hold a degree in CS/IT	48 (80%)	54 (90%)	52 (87%)
Computer Expertise			
Below or Somewhat Below Average	1 (2%)	1 (2%)	2 (3%)
Average	21 (35%)	22 (37%)	22 (37%)
Above or Somewhat Above Average	38 (63%)	37 (62%)	36 (60%)
Knowledge on Detecting Phishing			
Complete Novice	3 (5%)	1 (2%)	1 (2%)
Below Average	3 (5%)	4 (7%)	2 (3%)
Average	21 (35%)	31 (52%)	28 (47%)
Above Average	28 (47%)	20 (33%)	24 (40%)
Expert	5 (8%)	4 (7%)	5 (8%)
Years Spent Using Gmail			
Less than 4 years	13 (22%)	6 (10%)	11 (18%)
Greater than or equal to 4 years	47 (78%)	54 (90%)	49 (82%)

(86%) did not work in or have degrees in Computer Science, although 98% of participants across all surveys viewed their computer expertise as at or above average. Similarly, when asked about their skills detecting phishing, most participants claimed average or above average knowledge, with 8% on average claiming to be experts in detecting phishing. 85% of participants had been using Gmail for 4 or more years, meaning they were likely very familiar with its UI and accustomed to interacting with it. As a result, we present our findings as an upper bound on how average users will correctly absorb the information from Gmail's SII.

4.4 Analysis

Quantitative Analysis: We collected users' answers to multiple choice questions, multiple response (select all that apply) questions, open-ended question responses and their actions on our replica Gmail website. For our multiple response questions, we performed the test of proportions (z-test) to compare the responses following a prior study [39]. We cannot use a Chi-square test because the answers for our multiple response questions were not independently collected (i.e., a participant can choose multiple answers for each question). We used the Kruskal-Wallis test [90] and calibration curve [48] to compare confidence scores across groups.

We then use descriptive statistics to highlight the proportion of participants from each group that responded with specific answers. We present the proportion of participants that noticed 'via', selected specific behavior responses, and reported the proportion of participants with specific confidence scores. We use the statsmodels package in Python to conduct the analysis [69].

Qualitative Analysis: Two researchers on the team conducted iterative qualitative coding on the open-ended questions in the survey responses. (1) We asked participants to elaborate on the meaning of 'via' by asking "Please elaborate on what you think 'via' means". If participants reported that they did not understand the meaning of 'via', we asked them to guess: "What information do you think Gmail is trying to communicate by showing 'via' for this email? Please make your best guess and feel free to refer back to the email." (2) We asked participants to elaborate on the relationship between Gmail and 'via' by asking "Why do you think Gmail has chosen to display 'via' to users for certain emails?" (3) We asked participants to write down reasons "why they agree or disagree that Chase instructed the entity to send the email".

For each open-ended question, two researchers first conducted open coding to capture the major themes on 30% of the responses that were randomly selected. Then, the two researchers discussed and updated the codebook until an agreement about the themes was reached. We developed a codebook for each question to guide us in identifying the major themes for each condition. For example, participants were asked to explain why Gmail presented the 'via' in the email. One of the resulting codes for the Support group was "third party", which was used whenever a participant mentioned that Gmail provided the indicator to let them know the message was sent using a third party. Section A in the Appendix shows the code book and resulting themes.

After the training and codebook development, the two researchers coded all survey responses independently. After this initial coding, all codes reached acceptable inter-rater reliability (Cohen's Kappa above 0.7) [56]. The two researchers then talked through all instances where there was disagree-

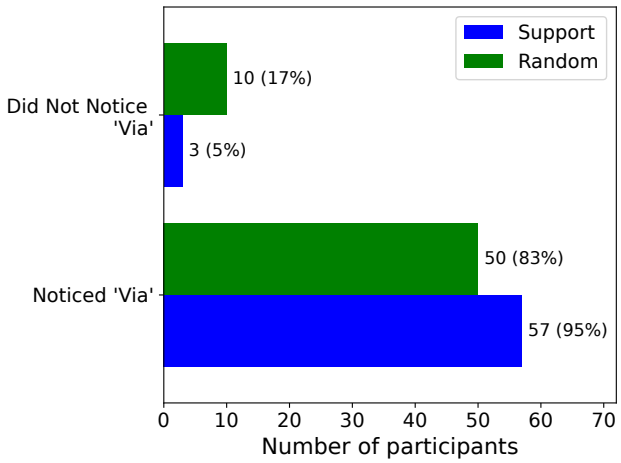


Figure 6: Number and percentage of participants that noticed 'via' in the study

ment and asked a third researcher to provide an opinion for judgment until a final decision was agreed upon.

5 Results

In this section, we present the qualitative and quantitative results of our study. We used this mixed-methods approach to understand participant behavior and indicator comprehension.

5.1 How do users respond to the presence of the 'via' indicator?

We explore the response of participants to the presence of the 'via' indicator by identifying (1) how many participants noticed 'via' when their attention is directed to the sender information section; (2) how many participants mentioned 'via' when determining the email sender; (3) how many participants checked the explanation of 'via' during the study.

The 'via' indicator was noticed by the majority of participants who were shown the indicator. Since security indicators are ineffective if they cannot capture users' attention [16], we asked participants if they noticed the 'via' indicator after they were asked to provide information about the email sender. Figure 6 shows that 89% of participants (n=120) noticed the 'via' indicator from the two groups that saw the 'via' indicator during the study. For the Support (n=60) and Random (n=60) groups respectively, 95% and 83% of the users in each group reported seeing the 'via' indicator during the study.[4]

While the notice rate is high, half of the participants believe they do not know the meaning of 'via'. After asking participants if they saw the 'via' indicator, we followed up by asking them if they knew the meaning of 'via'. Half of the participants (50%, n=120) reported not knowing the meaning of 'via' (22 in the Support group and 38 in the Random group). We hypothesize that this finding may be due to participant confidence and the limitations of 'via'. The indicator can only provide the origin domain for an email. It does not detect spoofing. Thus, instead of using 'via' as an aid to determine an email's origin, participants lean into their knowledge from prior experiences. Since they are confident about their ability to identify the email sender (the average confidence score is 4.61 and 4.70 out of 5.0 for the Support and Random group respectively), they might not care about the purpose or content of these indicators.

Given this low rate of understanding, we then examined the number of participants who clicked the indicator in our replica Gmail web browser, which provides an explanation of 'via'. Only 17 participants (4 in the Support group and 13 in the Random group) clicked the indicator while completing the study. We hypothesize that the low click-rate is mainly due to issues with indicator affordance — the indicator may not provide obvious visual cues that signal it can or should be clicked. Additionally, the fact that Prolific participants are motivated to complete the study quickly may have also contributed to the low click-rate.

Most participants did not mention 'via' when discussing the email sender. We asked participants to provide the email address of the sender, select how confident they were in their answer, and then discuss how they identified the information. Some users might not perceive the full difference between the email's true origin and its purported sender, but if the indicator works as intended, we expect experienced email users to acknowledge the via domain to some extent in their explanation, especially for the Random domain. Sadly, despite 89% of participants reportedly seeing 'via', and 50% of participants purportedly knowing the meaning of 'via', only 14% of participants mentioned the 'via' domain when explaining how they decided the sender of the email. Specifically, only eight participants (13%) in the Support group and nine participants (15%) in the Random group mentioned the 'via' domain to some extent in their answers. For example, P28 in the Support group specifically mentioned 'chase.com via chasesupport.com' and P52 in the Random group simply wrote *r1xaz.xyz* as their response. Lastly, only two participants, both from the Random group, raised concerns about identifying the email of the sender. For example, P40 in the Random group responded: *alerts@chase.com BUT there is a "via" thing after that that is new to me, and the explanation in the side window that pops up when I click on it about what "via" is, is not clear. If it weren't for that I'd be sure this came from Chase.com. But that "via" makes me wary*.

Additionally, despite most participants not mentioning the 'via' domain to some extent, the majority of participants were confident in their answer about the email sender when asked

[4] We note that prior work [24] has suggested that users can over-report their attention to security indicators. As such, our results represent the upper bound of the number of users who noticed the 'via' indicator.

Table 3: Confidence scores reported by participants when asked to provide the email address of the sender. The scale of 1 represents not confident and 5 represents very confident.

	Control n=60	**Support** n=60	**Random** n=60
5	50 (83%)	46 (76%)	46 (76%)
4	4 (7%)	8 (13%)	10 (17%)
3	1 (2%)	4 (7%)	4 (7%)
2	2 (3%)	1 (2%)	0
1	3 (5%)	1 (2%)	0

on a scale of 1 (not confident) to 5 (very confident): almost 80% of participants answered 5 (Table 3). After conducting a Kruskal-Wallis test, we found no statistically significant difference (p>.05) in confidence scores between each group. So while some users, like P40 in the Random group, were "wary", most participants were confident in their answers. This result suggests that, for most participants, 'via' is not a factor in identifying the origin of an email and does not lead to sender suspicion. The purpose of the indicator is to increase user awareness of email origin. If operating according to its purpose, more participants in the Random group would have confidently mentioned the 'via' domain (`r1xaz.xyz`) when discussing the email origin.

5.1.1 Calibration of Confidence

Following prior work [58, 59], we examine users' self-reported confidence against their actual performance in identifying the email address of the sender using a calibration curve [48], which is shown in Figure 7. The solid red line in Figure 7 represents perfect calibration, which is diagonal. When a user is perfectly calibrated, the probability of them mentioning 'via' in their answer is equal to their relative confidence in their answer (e.g., if a user is 80% confident in their answer, they would mention 'via' 80% of the time). Data points above the perfect calibration line correspond to users who are underconfident (e.g., if a user is 80% confident in their answer, they mention 'via' 90% of the time), while data points below the perfectly calibrated line correspond to users who are overconfident (e.g., if a user is 80% confident in their answer, they mention 'via' 70% of the time).

We derive the calibration curve for the Random group (orange dashed line) and the Support group (blue dotted line) by computing the rate of mentioning 'via' at each confidence level. We further convert the confidence level from 1 to 5 to a percentage scale (20% to 100%).

Overall, users are overconfident in both groups by a large margin. As past literature [65] has shown, confident users are less prone to change their online behavior and at a greater risk of being victimized. This result once again highlights that the 'via' indicator will not be effective in reducing unsafe behavior for users such as those in our study.

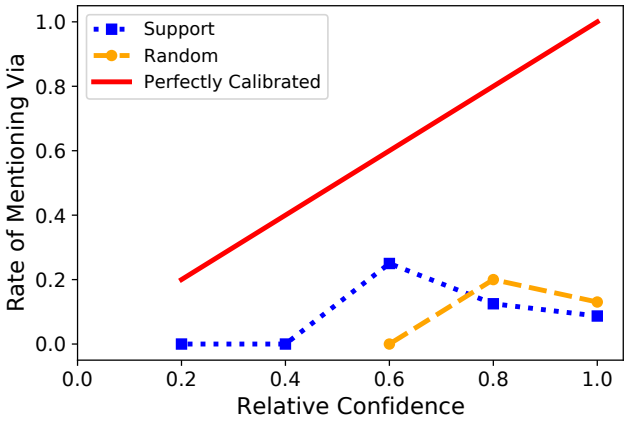


Figure 7: Calibration curve for identifying the email address of the sender.

5.2 How would users react to the email when the 'via' indicator is present?

To understand whether the 'via' indicator has an impact on participants' response to email messages that trigger 'via', we asked them what actions they would take after viewing the email shown in their group, as detailed in Section 4.2. Table 4 shows the number of participants from each group that selected the options provided. We compared the Random and Support group responses to the Control group responses using the test of proportions.

We did not observe statistically significant differences between the three groups (p>.05) for each action option. Most notably, over half the participants from each group selected that they would "Click on the view my summary button in the email". Also, none of the participants selected that they would "contact the bank in ways other than email", four participants selected they would "forward the email", and 38 (21%, n=180) participants selected that they would "delete the email", which are all actions Chase suggests people do if they receive a spoofed email [5]. This situation suggests that it is unlikely that the 'via' indicator encourages users to behave differently from when the indicator is not present.

5.3 How do users interpret the presence of the 'via' indicator?

Among the 120 participants that were shown the 'via' indicator, 47 participants (39%) marked that they knew its meaning. We examine if 'via' is effective at communicating the origin of the email to end users by asking all participants to explain or attempt to explain the purpose of the 'via' indicator.

Table 4: Number of participants from each group that selected the options provided

	Control n=60	**Support** n=60	**Random** n=60
Click button in email	39 (65%)	39 (65%)	35 (58%)
Archive the email	26 (43%)	30 (50%)	24 (40%)
Delete the email	13 (22%)	10 (17%)	15 (25%)
Other	2 (3%)	2 (3%)	5 (8%)
Reply by email	1 (2%)	2 (3%)	3 (5%)
Forward the email	0	1 (2%)	3 (5%)
Search Google	2 (3%)	1 (2%)	1 (2%)
Contact the bank	0	0	0

'via' means through. Many participants (21 in the Support group and 17 in the Random group) believed 'via' to mean through, as in this email was sent through a third party. For instance, P22 in the Support group wrote *"that the email came via an intermediary and not directly from chase.com*". In the Random group, P34 explained it as *"It's a return-path domain because the email was sent via a third party"*.

'via' indicates the sender. A significant amount of participants (18 in the Support group and 19 in the Random group) thought the 'via' domain indicated the true sender. For example, P21 in the Support group wrote *"That the e-mail was generated from a different domain name than the domain used in the actual e-mail sender @ address, I would assume like a mailer software that auto sent out the e-mails via a secondary support website."* and P6 in the Random group wrote *"That is the true origin of the email"*. P60 in the Random group took this explanation further and suggested that the email was forwarded, writing *"it could mean forwarded from i.e via but on reflection this is now likely a scam email phishing etc"*.

'via' indicates group association. When the target and 'via' domain include the bank brand name, participants associated the 'via' domain with the bank. Unique to the Support group, many participants (16) mention that 'via' means the email comes from an entity associated with Chase (e.g., Chase's support division). For example, P5 wrote that the email *"comes from a different department through the main company email"*. Additionally, two participants expanded on this idea, stating that 'via' indicated that an email was safe or authenticated. P25 from the Support group wrote *"That it's legitimately from Chase and not a scammer"*.

'via' encourages caution. When the 'via' domain does not include a brand name, participants explained that the presence of 'via' communicates a security risk. Many users (13) in the Random group explained that 'via' was being presented due to a scam or some other security risk that should be considered. While explaining the meaning of 'via', P59 in the Random group mentioned *"This [means] another website has been used to route the email. That's why I suspected it might be a phishing attempt"*. Additionally, P17 in the Random group wrote that the 'via' indicator *"possibly [means] that someone else is sending the email, looking at this more closely it appears like a scam"*.

This comparison suggests that the 'via' domain can influence users' interpretation of the 'via' indicator, especially when they try to guess the meaning. When the domain is shown as `chasesupport.com`, users are more willing to believe that this is related to Chase Bank, while the random domain triggered users' concerns about email safety.

5.4 What information do participants think Gmail is trying to communicate by showing 'via' in an email?

After participants explain what 'via' means, we then ask participants to reflect on why Gmail has chosen to display the 'via' indicator for the email they viewed in the study. In this section, we show how prompting participants to consider Gmail's perspective changes their interpretation.

In contrast to participants' explanation of 'via', security is one of the common reasons why many participants think Gmail has chosen to display 'via'. Of the 120 participants that saw 'via' in the study, 49 (41%) participants (27 from Support and 22 from Random) think Gmail chose to display 'via' to warn them of phishing or email legitimacy. For example, users explained that 'via' is displayed *"to make sure the email is not fraudulent"*[P18, Support]. Some users specifically mentioned security issues like phishing email or scam email: *"So people know the true email it came from cause it could actually be a scammer"*[P44, Random].

In addition to security, many participants believe Gmail uses 'via' to provide additional information about the sender. Many of the participants from both groups, 51 (43%) participants (24 from Support and 27 from Random) think 'via' is displayed to provide additional information. Some participants expressed this belief, writing that Gmail wants the user to know the email was outsourced to a third party or was not sent directly from Chase. Others explained it as Gmail wanting to provide additional transparency to the email. For example, P32 in the Support group elaborated that *"Gmail has chosen to display via to show where the email has come from if the recipient wants to check out the website."* However, some participants also connected this transparency to authenticating the sender — *"I think Gmail is adding it to certain emails to add authenticity"*[P9, Random].

5.5 What are users' perceptions of the relationship between the 'via' domain and `chase.com`?

After we asked users to think about what 'via' means and why Gmail chose to display it, we asked participants if they thought `chase.com` used the 'via' domain to send them the email and

explain their reasoning. The 'via' indicator only implies that the actual sender (which used the 'via' domain) of an email is different than the purported sender (with domain name chase.com), and the indicator was not intended to signal a relationship between the two domains.

Most participants from the Support (73%, n=60) and Random group (53%, n=60) believed that chase.com used the 'via' domain to send the email. Many participants (62%, n=120) believed that Chase Bank or chase.com used the 'via' domain to send the email because they believed the 'via' domain was the sending service, the domain chasesupport.com appears to be a part of the Chase business, or because this order of events matches their explanation of 'via'.

Some participants believed that chasesupport.com was a part of the Chase Bank business. Unique to the Support group, some participants (12) explicitly signaled the relationship between the domain chasesupport.com and Chase. For example, P18 described that *"even though both emails are from the same company, one division chase.com asked or used information from chasesupport.com."* This also includes participants who believe the email was initiated by Chase (5 participants) or that Google had verified the email (2 participants). Others (4 participants) in the group believed the two domains were associated, but that chasesupport.com instructed chase.com to send the email. An example of this perspective is from P21 in the Support group who wrote, *"The way that I think the 'via' works would, in my mind, mean that the chasesupport site auto-generated the e-mail and instructed the chase.com address to send the e-mail, not the other way around."* This result indicates the impact brand name has on user interpretations. We asked participants to explain the purpose of 'via' in their own words, from the perspective of Gmail, and then in relation to the target domain. In every scenario, multiple participants from the Support group viewed chasesupport.com as an authentic domain associated with Chase Bank.

Overall, only a small portion (6 in the Support group and 17 in the Random group) of participants were able to determine that chase.com did not use the 'via' domain to send the email and expressed some level of security concerns. Some of these users specifically mentioned the possibility of email being falsified and others raised some level of suspicion. P24 in the Support group correctly stated that *"The email was sent from chasesupport.com. chase.com didn't 'use' or 'instruct' anything. chasesupport.com sent the email"*. However, we note that even though this participant was able to correctly interpret the relationship between chase.com and the 'via' domain, they indicated that they would "Click on the view my summary button in the email" in the beginning of the study.

In fact, this apparent contradiction is not rare: after going through the questions, some participants realized what 'via' was communicating and expressed a new opinion of their previous actions. When discussing this question P8, in the Random group, wrote *"Had I seen the 'via' and the scary lookin' link, definitely would've just flagged this email, but it was sort of inconspicuous."* When asked how they would respond to the email, P8 selected: "Keep, save or archive the email"; "Click on the view my summary button in the email"; and "Delete the email". Thus, after taking time to reflect on 'via' from multiple perspectives, this participant was able to change their original decision. In fact, a non-negligible amount (5 in the Support group and 10 in the Random group) of participants expressed security issues after saying they would "Click on the view my summary button in the email" in response to the email earlier in the study. This result suggests that the indicator can be interpreted but is unlikely to nudge new behavior during the real-time decision-making process.

6 Limitations

The results of our survey are limited by the chosen scenario, the use of self-reported data, and participant demographics.

The Support, Random and Control group participants were all shown an email message that was supposedly from chase.com (which belongs to Chase Bank). As such, users' prior experience with Chase and prior exposure to email from Chase may have an impact on their responses. We also note that chase.com has a strong DMARC policy, and the example we show here is not representative of what might actually happen when chase.com is spoofed. However, research suggests that this type of spoofing with email forwarding can be done for other brands [51], thus we use chase.com to represent that possibility in the study. We chose a well-known brand name to investigate participant reactions when the target domain and 'via' domain include the brand name.

Next, while we strive to mimic users' real experience with Gmail, participants may act differently in our study compared to what they might do when using their personal email account. However, unlike other studies in this area, we focus on the differences in behavior selections due to indicator presence instead of focusing on one specific behavior under different email message conditions. We also recognize that users have access to the email client throughout the study, which may impact some of the self-reported results (e.g., "do you remember seeing 'via'?"). However, we believe users' response to this question does not negate their interpretation of the indicator's purpose. Design guidelines advise that warnings and indicators be noticeable and easy to interpret and, thus, should not require prior experience [47].

Lastly, our results may not generalize to the US and UK populations. Prolific users are more knowledgeable about security and have more confidence about that knowledge [76]. Due to this potential bias, we view our results as a reflection of how technically skilled individuals perceive the 'via' indicator.

7 Discussion

There are many individual challenges that, together, undermine a user's ability to effectively incorporate Gmail's "via" indicator into their decision making.

Among these are the general challenges associated with passive indicators (e.g., as highlighted in the context of phishing [22, 93]). Our work similarly documents that passive indicators are not able to prompt users to make safe decisions about email origin. This, in part, is because users often consider security as a secondary task [16] and rarely invest time and attention engaging in questions of security [32, 67]. In our work, we show that the presence of the 'via' indicator has little impact on whether or not users click on an embedded link. This result holds true even when we intentionally draw users' attention to the sender information section and even though the majority of users report that they noticed the 'via' indicator. In practice (i.e., without such prompting), it is likely that many users will overlook the 'via' indicator entirely: 'via' has a light gray color, is the same size as the rest of the header text, and is semantically vague without clearly conveying any notion of risk. The 'via' indicator could be made more noticeable if it were changed to a color that contrasts more with the surrounding text, and if the word used for the indicator were more related to its intended security purpose.

While this study focused on the 'via' indicator in the desktop version of Gmail, we also point out that the mobile Gmail app has no 'via' indicator at all. To get the same information provided by the 'via' on desktop, the mobile user has to open the collapsed box of sender details, and then click on the 'View security details' link to find the 'Mailed by' field. Using the phrase 'mailed by' instead of 'via' is an improvement, as it more accurately conveys the purpose of this indicator. However, the fact that this information is hidden behind two easy-to-miss interactions means that the chance of users finding this information is even lower.

Another major obstacle that hinders the success of 'via' is that it relies upon the ability of users to correctly interpret its meaning and the domain displayed after it. In our work, we show that the domain displayed after 'via' heavily influences users' interpretation of the indicator and their perception of the security risk. This once again highlights the potential issue of relying on users' computer knowledge. On one hand, past literature has consistently suggested that users cannot reliably determine the legitimacy of a domain [3, 4]. On the other hand, sometimes it is naturally a difficult undertaking to decide which domain names are connected with a specific organization [4].

A third issue, which is also common among security warnings, is the need for clearness in explanation. Participants in our study have indicated difficulty in comprehending the explanation of 'via' — if they were even able to find the explanation for the 'via' indicator in the first place. Indeed, upon carefully examining the current explanation of 'via', it does not convey the potential security risks in a straightforward way, and contains jargon that can be hard for users to understand. Given that the majority of Gmail users will not be familiar with DMARC policies or domain names, the current explanation does not fulfill its goal of explaining what the 'via' indicator means to the user. To improve both comprehension and safe behavior, the explanation should be updated with a more approachable explanation that highlights the potential security risks at the beginning, leaving the more technical details for the end so advanced users can still find it.

Last but not least, the current design introduces a new layer of complexity for users to determine spoofed email messages. Past papers have suggested that checking if the sender email address and organization are the same is a good approach to determine whether an email is legitimate [58, 100]. However, in the case of 'via', it is possible for the sender email and name to match, while also having a different domain as the 'via' domain. This situation makes it even more challenging for users to determine whether an email message is legitimate, as they are used to having to check that only two pieces of information match. This additional piece of information means users have to learn how to process more indicators, but once they learn how to do so, it can enable them to make safer decisions regarding which email messages to interact with. Gmail is one of the more secure email clients in this sense, as most clients do not display the 'via' information at all. These other clients without indicators prevent users from becoming too overwhelmed by information and warnings, but at the same time possibly expose them to greater risk of phishing or other unsafe situations. There is no an easy answer for this dilemma, since it involves a carefully balancing act of enabling users to make safe and informed decisions without succumbing to warning fatigue.

8 Conclusion

In this paper, we present a first analysis of the effectiveness and comprehensibility of Gmail's 'via' indicator. We conduct a survey to evaluate whether users notice the 'via' indicator, whether they understand the meaning of the 'via' indicator, and how their understanding affects what action users take with the email. We find that the majority of participants notice the 'via' indicator, but still proceed with unsafe behavior due to misunderstandings of what the indicator represents. Additionally, the understanding of what the 'via' indicator represents is heavily influenced by how familiar users are with the 'via' domain. The use of a more familiar and seemingly trustworthy domain thwarted the 'via' indicator's intended goal of conveying potential security concerns. These findings highlight the shortcomings of current passive security indicator design, and emphasize the need for indicators that users will notice and understand while still providing salient security information.

Acknowledgments

We thank our anonymous reviewers for their constructive feedback that helps make this paper better. We thank Kristen Vaccaro and Mary Anne Smart for their help with designing the study. We thank Cindy Moore and Jennifer Folkestad for their operational support. Funding for this work was provided in part by National Science Foundation grant CNS-2152644, the UCSD CSE Postdoctoral Fellows program, the Irwin Mark and Joan Klein Jacobs Chair in Information and Computer Science.

References

[1] Saeed Abu-Nimeh, Dario Nappa, Xinlei Wang, and Suku Nair. A Comparison of Machine Learning Techniques for Phishing Detection. In *Proceedings of the Anti-phishing Working Groups 2nd Annual Ecrime Researchers Summit*, pages 60–69, 2007.

[2] Devdatta Akhawe and Adrienne Porter Felt. Alice in Warningland: A Large-Scale Field Study of Browser Security Warning Effectiveness. In *Proceedings of the 22nd USENIX Security Symposium*, 2013.

[3] Sara Albakry, Kami Vaniea, and Maria K. Wolters. What Is This Url's Destination? Empirical Evaluation of Users' Url Reading. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pages 1–12, 2020.

[4] Kholoud Althobaiti, Nicole Meng, and Kami Vaniea. I Don't Need an Expert! Making Url Phishing Features Human Comprehensible. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pages 1–17, 2021.

[5] Chase Bank. Frequently Asked Questions: Fraud, 01 2023. https://www.chase.com/digital/resources/privacy-security/questions/fraud.

[6] Maxim Baryshevtsev and Joseph McGlynn. Persuasive Appeals Predict Credibility Judgments of Phishing Messages. *Cyberpsychology, Behavior, and Social Networking*, 23(5):297–302, 2020.

[7] Simon Bell and Peter Komisarczuk. An Analysis of Phishing Blacklists: Google Safe Browsing, Openphish, and Phishtank. In *Proceedings of the 2020 Australasian Computer Science Week Multiconference*, pages 1–11, 2020.

[8] Zinaida Benenson, Freya Gassmann, and Robert Landwirth. Unpacking Spear Phishing Susceptibility. In *Proceedings of the 2017 International Conference on Financial Cryptography and Data Security*, pages 610–627. Springer, 2017.

[9] Mark Blythe, Helen Petrie, and John A. Clark. F for Fake: Four Studies on How We Fall for Phish. In *Proceedings of the 2011 CHI Conference on Human Factors in Computing Systems*, pages 3469–3478, 2011.

[10] Paolo Buono, Giuseppe Desolda, Francesco Greco, and Antonio Piccinno. Let Warnings Interrupt the Interaction and Explain: Designing and Evaluating Phishing Email Warnings. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*, pages 1–6, 2023.

[11] Deanna D. Caputo, Shari Lawrence Pfleeger, Jesse D. Freeman, and M. Eric Johnson. Going Spear Phishing: Exploring Embedded Training and Awareness. *IEEE Security & Privacy*, 2013.

[12] Jianjun Chen, Vern Paxson, and Jian Jiang. Composition Kills: A Case Study of Email Sender Authentication. In *Proceedings of the 29th USENIX Security Symposium*, pages 2183–2199, 2020.

[13] Debra L. Cook, Vijay K. Gurbani, and Michael Daniluk. Phishwish: A Simple and Stateless Phishing Filter. *Security and Communication Networks*, 2009.

[14] Marco De Bona and Federica Paci. A Real World Study on Employees' Susceptibility to Phishing Attacks. In *Proceedings of the 15th International Conference on Availability, Reliability and Security*, pages 1–10, 2020.

[15] Rachna Dhamija and J. D. Tygar. The Battle Against Phishing: Dynamic Security Skins. In *Proceedings of the 2005 Symposium on Usable Privacy and Security*, pages 77–88, 2005.

[16] Rachna Dhamija, J. D. Tygar, and Marti Hearst. Why Phishing Works. In *Proceedings of the 2006 CHI Conference on Human Factors in Computing Systems*, pages 581–590, 2006.

[17] Alejandra Diaz, Alan T. Sherman, and Anupam Joshi. Phishing in an Academic Community: A Study of User Susceptibility and Behavior. *Cryptologia*, 2020.

[18] Julie S. Downs, Mandy Holbrook, and Lorrie Faith Cranor. Behavioral Response to Phishing Risk. In *Proceedings of the Anti-Phishing Working Groups 2nd Annual eCrime Researchers Summit*, pages 37–44, 2007.

[19] Julie S. Downs, Mandy B. Holbrook, and Lorrie Faith Cranor. Decision Strategies and Susceptibility to Phishing. In *Proceedings of the Second Symposium on Usable Privacy and Security*, pages 79–90, 2006.

[20] Sevtap Duman, Kubra Kalkan-Cakmakci, Manuel Egele, William Robertson, and Engin Kirda. Emailprofiler: Spearphishing Filtering With Header and Stylometric Features of Emails. In *Proceedings of the 2016 IEEE Annual Computer Software and Applications Conference*, pages 408–416, 2016.

[21] EasyDMARC. Email Forwarding and DMARC DKIM SPF, 05 2022. https://easydmarc.com/blog/email-forwarding-and-dmarc-dkim-spf/.

[22] Serge Egelman, Lorrie Faith Cranor, and Jason Hong. You've Been Warned: An Empirical Study of the Effectiveness of Web Browser Phishing Warnings. In *Proceedings of the 2009 CHI Conference on Human Factors in Computing Systems*, pages 1065–1074, 2008.

[23] Franz Faul, Edgar Erdfelder, Axel Buchner, and Albert-Georg Lang. Statistical Power Analyses Using G* Power 3.1: Tests for Correlation and Regression Analyses. *Behavior research methods*, 41(4):1149–1160, 2009.

[24] Adrienne Porter Felt, Elizabeth Ha, Serge Egelman, Ariel Haney, Erika Chin, and David Wagner. Android Permissions: User Attention, Comprehension, and Behavior. In *Proceedings of the Eighth Symposium on Usable Privacy and Security*, pages 1–14, 2012.

[25] Ian Fette, Norman Sadeh, and Anthony Tomasic. Learning to Detect Phishing Emails. In *Proceedings of the 16th International Conference on World Wide Web*, pages 649–656, 2007.

[26] Anjuli Franz, Verena Zimmermann, Gregor Albrecht, Katrin Hartwig, Christian Reuter, Alexander Benlian, Joachim Vogt, et al. SOK: Still Plenty of Phish in the Sea-a Taxonomy of User-Oriented Phishing Interventions and Avenues for Future Research. In *Proceedings of Seventeenth Symposium on Usable Privacy and Security*, pages 339–358, 2021.

[27] Sujata Garera, Niels Provos, Monica Chew, and Aviel D. Rubin. A Framework for Detection and Measurement of Phishing Attacks. In *Proceedings of the 2007 ACM Workshop on Recurring Malcode*, pages 1–8, 2007.

[28] Google. Google Safe Browsing, 01 2023. https://safebrowsing.google.com/.

[29] Ayako Akiyama Hasegawa, Naomi Yamashita, Mitsuaki Akiyama, and Tatsuya Mori. Why They Ignore English Emails: The Challenges of Non-Native Speakers in Identifying Phishing Emails. In *Proceedings of Seventeenth Symposium on Usable Privacy and Security*, pages 319–338, 2021.

[30] Farkhondeh Hassandoust, Harminder Singh, and Jocelyn Williams. The Role of Contextualization in Individuals' Vulnerability to Phishing Attempts. *Australasian Journal of Information Systems*, 2020.

[31] Ryan Heartfield and George Loukas. A Taxonomy of Attacks and a Survey of Defence Mechanisms for Semantic Social Engineering Attacks. *ACM Computing Surveys*, 2015.

[32] Cormac Herley. So Long, and No Thanks for the Externalities: The Rational Rejection of Security Advice by Users. In *Proceedings of the 2009 Workshop on New Security Paradigms Workshop*, pages 133–144, 2009.

[33] Amir Herzberg and Ahmad Gbara. Trustbar: Protecting (Even Naive) Web Users From Spoofing and Phishing Attacks. Technical report, 2004. http://eprint.iacr.org/2004/155.

[34] Grant Ho, Aashish Sharma, Mobin Javed, Vern Paxson, and David Wagner. Detecting Credential Spearphishing in Enterprise Settings. In *Proceedings of the 2017 USENIX Security Symposium*, pages 469–485, 2017.

[35] Hang Hu and Gang Wang. End-to-End Measurements of Email Spoofing Attacks. In *Proceedings of the 27th USENIX Security Symposium*, pages 1095–1112, 2018.

[36] Collin Jackson, Daniel R. Simon, Desney S. Tan, and Adam Barth. An Evaluation of Extended Validation and Picture-in-Picture Phishing Attacks. In *Proceedings of the 2007 International Conference on Financial Cryptography and Data Security*, pages 281–293. Springer, 2007.

[37] Daniel Jampen, Gürkan Gür, Thomas Sutter, and Bernhard Tellenbach. Don't Click: Towards an Effective Anti-Phishing Training. A Comparative Literature Review. *Human-centric Computing and Information Sciences*, 2020.

[38] Yogesh Joshi, Samir Saklikar, Debabrata Das, and Subir Saha. Phishguard: A Browser Plug-in for Protection From Phishing. In *Proceedings of the 2008 International Conference on Internet Multimedia Services Architecture and Applications*, pages 1–6. IEEE, 2008.

[39] Smirity Kaushik, Yaxing Yao, Pierre Dewitte, and Yang Wang. "How I Know for Sure": People's Perspectives on Solely Automated Decision-Making (SADM). In *Proceedings of the Seventeenth Symposium on Usable Privacy and Security*, pages 159–180, 2021.

[40] Mahmoud Khonji, Youssef Iraqi, and Andrew Jones. Mitigation of Spear Phishing Attacks: A Content-Based Authorship Identification Framework. In *Proceedings of the 2011 International Conference for Internet Technology and Secured Transactions*, 2011.

[41] Ponnurangam Kumaraguru, Justin Cranshaw, Alessandro Acquisti, Lorrie Cranor, Jason Hong, Mary Ann Blair, and Theodore Pham. School of Phish: A Real-World Evaluation of Anti-Phishing Training. In *Proceedings of the Fifth Symposium on Usable Privacy and Security*, pages 1–12, 2009.

[42] Ponnurangam Kumaraguru, Yong Rhee, Alessandro Acquisti, Lorrie Faith Cranor, Jason Hong, and Elizabeth Nunge. Protecting People From Phishing: The Design and Evaluation of an Embedded Training Email System. In *Proceedings of the 2007 CHI Conference on Human Factors in Computing Systems*, pages 905–914, 2007.

[43] Ponnurangam Kumaraguru, Yong Rhee, Steve Sheng, Sharique Hasan, Alessandro Acquisti, Lorrie Faith Cranor, and Jason Hong. Getting Users to Pay Attention to Anti-Phishing Education: Evaluation of Retention and Transfer. In *Proceedings of the Anti-phishing Working Groups 2nd Annual eCrime Researchers Summit*, 2007.

[44] Ponnurangam Kumaraguru, Steve Sheng, Alessandro Acquisti, Lorrie Faith Cranor, and Jason Hong. Lessons From a Real World Evaluation of Anti-Phishing Training. In *Proceedings of the Anti-phishing Working Groups 3rd Annual eCrime Researchers Summit*, 2008.

[45] Ponnurangam Kumaraguru, Steve Sheng, Alessandro Acquisti, Lorrie Faith Cranor, and Jason Hong. Teaching Johnny Not to Fall for Phish. *ACM Transactions on Internet Technology*, 2010.

[46] Elmer Lastdrager, Inés Carvajal Gallardo, Pieter Hartel, and Marianne Junger. How Effective Is Anti-Phishing Training for Children? In *Proceedings of the Thirteenth Symposium on Usable Privacy and Security*, pages 229–239, 2017.

[47] Kenneth R. Laughery and Michael S. Wogalter. Designing Effective Warnings. *Reviews of Human Factors and Ergonomics*, 2006.

[48] Sarah Lichtenstein, Baruch Fischhoff, and Lawrence D. Phillips. Calibration of Probabilities: The State of the Art. In *Proceedings of the Fifth Research Conference on Subjective Probability, Utility, and Decision Making*, pages 275–324, 1977.

[49] Eric Lin, Saul Greenberg, Eileah Trotter, David Ma, and John Aycock. Does Domain Highlighting Help People Identify Phishing Sites? In *Proceedings of the 2011 CHI Conference on Human Factors in Computing Systems*, pages 2075–2084, 2011.

[50] Tian Lin, Daniel E. Capecci, Donovan M. Ellis, Harold A. Rocha, Sandeep Dommaraju, Daniela S. Oliveira, and Natalie C. Ebner. Susceptibility to Spear-Phishing Emails: Effects of Internet User Demographics and Email Content. *ACM Transactions on Computer-Human Interaction*, 2019.

[51] Enze Liu, Gautam Akiwate, Mattijs Jonker, Ariana Mirian, Grant Ho, Geoffrey M. Voelker, and Stefan Savage. Forward Pass: On the Security Implications of Email Forwarding Mechanism and Policy. In *Proceedings of the 8th IEEE European Symposium on Security and Privacy*, 2023.

[52] Enze Liu, Gautam Akiwate, Mattijs Jonker, Ariana Mirian, Stefan Savage, and Geoffrey M. Voelker. Who's Got Your Mail? Characterizing Mail Service Provider Usage. In *Proceedings of the 21st ACM Internet Measurement Conference*, pages 122–136, 2021.

[53] Gang Liu, Guang Xiang, Bryan A. Pendleton, Jason I. Hong, and Wenyin Liu. Smartening the Crowds: Computational Techniques for Improving Human Verification to Fight Phishing Scams. In *Proceedings of the Seventh Symposium on Usable Privacy and Security*, pages 1–13, 2011.

[54] Samuel Marchal, Giovanni Armano, Tommi Gröndahl, Kalle Saari, Nidhi Singh, and N Asokan. Off-the-Hook: An Efficient and Usable Client-Side Phishing Prevention Application. *IEEE Transactions on Computers*, 2017.

[55] Samuel Marchal, Kalle Saari, Nidhi Singh, and N Asokan. Know Your Phish: Novel Techniques for Detecting Phishing Sites and Their Targets. In *Proceedings of the 36th International Conference on Distributed Computing Systems*, pages 323–333, 2016.

[56] Mary L. McHugh. Interrater Reliability: The Kappa Statistic. *Biochemia medica*, 2012.

[57] Gregory D. Moody, Dennis F. Galletta, and Brian Kimball Dunn. Which Phish Get Caught? An Exploratory Study of Individuals' Susceptibility to Phishing. *European Journal of Information Systems*, 2017.

[58] James Nicholson, Lynne Coventry, and Pam Briggs. Can We Fight Social Engineering Attacks by Social Means? Assessing Social Salience as a Means to Improve Phish Detection. In *Proceedings of the Thirteenth Symposium on Usable Privacy and Security*, pages 285–298, 2017.

[59] James Nicholson, Yousra Javed, Matt Dixon, Lynne Coventry, Opeyemi Dele Ajayi, and Philip Anderson. Investigating Teenagers' Ability to Detect Phishing Messages. In *Proceedings of the 2020 IEEE European Symposium on Security and Privacy Workshops*, 2020.

[60] Daniela Oliveira, Harold Rocha, Huizi Yang, Donovan Ellis, Sandeep Dommaraju, Melis Muradoglu, Devon Weir, Adam Soliman, Tian Lin, and Natalie Ebner. Dissecting Spear Phishing Emails for Older vs Young Adults: On the Interplay of Weapons of Influence and Life Domains in Predicting Susceptibility to Phishing. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pages 6412–6424, 2017.

[61] Justin Petelka, Yixin Zou, and Florian Schaub. Put Your Warning Where Your Link Is: Improving and Evaluating Email Phishing Warnings. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pages 1–15, 2019.

[62] PhishTank. Phishtank, 01 2023. https://phishtank.org/.

[63] ProofPoint. Phishing Protection, 01 2023. https://www.proofpoint.com/us/solutions/protect-against-phishing.

[64] Benjamin Reinheimer, Lukas Aldag, Peter Mayer, Mattia Mossano, Reyhan Duezguen, Bettina Lofthouse, Tatiana Von Landesberger, and Melanie Volkamer. An Investigation of Phishing Awareness and Education Over Time: When and How to Best Remind Users. In *Proceedings of the Sixteenth Symposium on Usable Privacy and Security*, pages 259–284, 2020.

[65] Markus Riek, Rainer Bohme, and Tyler Moore. Measuring the Influence of Perceived Cybercrime Risk on Online Service Avoidance. *IEEE Transactions on Dependable and Secure Computing*, 2015.

[66] Stefan A. Robila and James W. Ragucci. Don't Be a Phish: Steps in User Education. *ACM SIGCSE Bulletin*, 2006.

[67] Martina Angela Sasse, Sacha Brostoff, and Dirk Weirich. Transforming the 'Weakest Link'—a Human/Computer Interaction Approach to Usable and Effective Security. *BT Technology Journal*, 2001.

[68] Katharina Schiller, Florian Adamsky, and Zinaida Benenson. Towards an Empirical Study to Determine the Effectiveness of Support Systems Against E-Mail Phishing Attacks. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*, pages 1–15, 2023.

[69] Skipper Seabold and Josef Perktold. Statsmodels: Econometric and Statistical Modeling With Python. In *Proceedings of the 9th Python in Science Conference*, 2010.

[70] Kaiwen Shen, Chuhan Wang, Minglei Guo, Xiaofeng Zheng, Chaoyi Lu, Baojun Liu, Yuxuan Zhao, Shuang Hao, Haixin Duan, Qingfeng Pan, et al. Weak Links in Authentication Chains: A Large-Scale Analysis of Email Sender Spoofing Attacks. In *Proceedings of the 2021 USENIX Security Symposium*, 2021.

[71] Steve Sheng, Mandy Holbrook, Ponnurangam Kumaraguru, Lorrie Faith Cranor, and Julie Downs. Who Falls for Phish? A Demographic Analysis of Phishing Susceptibility and Effectiveness of Interventions. In *Proceedings of the 2010 CHI conference on Human Factors in Computing Systems*, pages 373–382, 2010.

[72] Steve Sheng, Bryant Magnien, Ponnurangam Kumaraguru, Alessandro Acquisti, Lorrie Faith Cranor, Jason Hong, and Elizabeth Nunge. Anti-Phishing Phil: The Design and Evaluation of a Game That Teaches People Not to Fall for Phish. In *Proceedings of the Third Symposium on Usable Privacy and Security*, pages 88–99, 2007.

[73] Kuldeep Singh, Palvi Aggarwal, Prashanth Rajivan, and Cleotilde Gonzalez. Training to Detect Phishing Emails: Effects of the Frequency of Experienced Phishing Emails. In *Proceedings of the 2019 Human Factors and Ergonomics Society Annual Meeting*, 2019.

[74] Eric Spero and Robert Biddle. Out of Sight, Out of Mind: Ui Design and the Inhibition of Mental Models of Security. In *Proceedings of the 2020 New Security Paradigms Workshop*, pages 127–143, 2020.

[75] Gianluca Stringhini and Olivier Thonnard. That Ain't You: Blocking Spearphishing Through Behavioral Modelling. In *Proceedings of the 2015 International Conference on Detection of Intrusions and Malware, and Vulnerability Assessment*, pages 78–97, 2015.

[76] Jenny Tang, Eleanor Birrell, and Ada Lerner. Replication: How Well Do My Results Generalize Now? The External Validity of Online Privacy and Security Surveys. In *Proceedings of the Eighteenth Symposium on Usable Privacy and Security*, pages 367–385, 2022.

[77] NCRLY Teraguchi and John C. Mitchell. Client-Side Defense Against Web-Based Identity Theft. *Computer Science Department, Stanford University*, 2004.

[78] Arun Vishwanath, Brynne Harrison, and Yu Jie Ng. Suspicion, Cognition, and Automaticity Model of Phishing Susceptibility. *Communication Research*, pages 1146–1166, 2018.

[79] Arun Vishwanath, Tejaswini Herath, Rui Chen, Jingguo Wang, and H. Raghav Rao. Why Do People Get Phished? Testing Individual Differences in Phishing Vulnerability Within an Integrated, Information Processing Model. *Decision Support Systems*, 2011.

[80] Melanie Volkamer, Karen Renaud, and Paul Gerber. Spot the Phish by Checking the Pruned URL. *Information & Computer Security*, 2016.

[81] Melanie Volkamer, Karen Renaud, Benjamin Reinheimer, and Alexandra Kunz. User Experiences of Torpedo: Tooltip-Powered Phishing Email Detection. *Computers & Security*, 2017.

[82] Chenkai Wang and Gang Wang. Revisiting Email Forwarding Security Under the Authenticated Received Chain Protocol. In *Proceedings of the 2022 ACM Web Conference*, pages 681–689, 2022.

[83] Ge Wang, He Liu, Sebastian Becerra, Kai Wang, Serge J Belongie, Hovav Shacham, and Stefan Savage. Verilogo: Proactive Phishing Detection via Logo Recognition. *Department of Computer Science and Engineering, University of California*, 2011.

[84] Rick Wash. How Experts Detect Phishing Scam Emails. *Proceedings of the ACM on Human-Computer Interaction*, 2020.

[85] Rick Wash and Molly M Cooper. Who Provides Phishing Training? Facts, Stories, and People Like Me. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pages 1–12, 2018.

[86] Rick Wash, Norbert Nthala, and Emilee Rader. Knowledge and Capabilities That Non-Expert Users Bring to Phishing Detection. In *Proceedings of Seventeenth Symposium on Usable Privacy and Security*, 2021.

[87] Zikai Alex Wen, Zhiqiu Lin, Rowena Chen, and Erik Andersen. What. Hack: Engaging Anti-Phishing Training Through a Role-Playing Phishing Simulation Game. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 2019.

[88] Liu Wenyin, Ning Fang, Xiaojun Quan, Bite Qiu, and Gang Liu. Discovering Phishing Target Based on Semantic Link Network. *Future Generation Computer Systems*, 2010.

[89] Colin Whittaker, Brian Ryner, and Marria Nazif. Large-Scale Automatic Classification of Phishing Pages. In *Proceedings of the 2010 Network and Distributed System Security Symposium*, 2010.

[90] Wikipedia. Kruskal-Wallis OneWay Analysis of Variance, 06 2023. https://en.wikipedia.org/wiki/Kruskal%E2%80%93Wallis_one-way_analysis_of_variance.

[91] Emma J Williams and Danielle Polage. How Persuasive Is Phishing Email? The Role of Authentic Design, Influence and Current Events in Email Judgements. *Behaviour and Information Technology*, 2019.

[92] Ryan T. Wright and Kent Marett. The Influence of Experiential and Dispositional Factors in Phishing: An Empirical Investigation of the Deceived. *Journal of Management Information Systems*, 2010.

[93] Min Wu, Robert C. Miller, and Simson L. Garfinkel. Do Security Toolbars Actually Prevent Phishing Attacks? In *Proceedings of the 2006 CHI Conference on Human Factors in Computing Systems*, 2006.

[94] Guang Xiang, Jason Hong, Carolyn P. Rose, and Lorrie Cranor. Cantina+ a Feature-Rich Machine Learning Framework for Detecting Phishing Web Sites. *ACM Transactions on Information and System Security*, 2011.

[95] Weining Yang, Aiping Xiong, Jing Chen, Robert W. Proctor, and Ninghui Li. Use of Phishing Training to Improve Security Warning Compliance: Evidence From a Field Experiment. In *Proceedings of the 2017 Hot Topics in Science of Security: Symposium and Bootcamp*, pages 52–61, 2017.

[96] Ka-Ping Yee and Kragen Sitaker. Passpet: Convenient Password Management and Phishing Protection. In *Proceedings of the Second Symposium on Usable Privacy and Security*, pages 32–43, 2006.

[97] Yaman Yu, Saidivya Ashok, Smirity Kaushi, Yang Wang, and Gang Wang. Design and Evaluation of Inclusive Email Security Indicators for People With Visual Impairments. In *Proceedings of the 2023 IEEE Symposium on Security and Privacy*, 2023.

[98] Yue Zhang, Serge Egelman, Lorrie Cranor, and Jason Hong. Phinding Phish: Evaluating Anti-Phishing Tools. *Carnegie Mellon University*, 2007.

[99] Yue Zhang, Jason I. Hong, and Lorrie F. Cranor. Cantina: A Content-Based Approach to Detecting Phishing Web Sites. In *Proceedings of the 16th International Conference on World Wide Web*, 2007.

[100] Sarah Zheng and Ingolf Becker. Presenting Suspicious Details in User-Facing E-Mail Headers Does Not Improve Phishing Detection. In *Proceedings of the Eighteenth Symposium on Usable Privacy and Security*, pages 253–271, 2022.

A Appendix

A.1 Survey

Prescreening Questions

1. Are you currently a Gmail user? [Yes/No]
2. How long have you been using Gmail? [Less than a year - Four or more years]

Survey Questions

1. Please provide the name of the person or entity that sent the email?
2. How confident are you that your answer is correct? [1-5]
3. Please provide the email address of the person or entity that sent the email?
4. How confident are you that your answer is correct? [1-5]
5. How did you decide who the sender of the email was?
6. Do you remember seeing something similar to what is highlighted in the picture shown below during the study? (yes & I know what it means/I don't know what it means or no)
7. Please elaborate on what you think 'via' means
8. What information do you think Gmail is trying to communicate by showing 'via' for this email? Please make your best guess and feel free to refer back to the email.
9. Why do you think Gmail has chosen to display 'via' to users for certain emails? Please make your best guess and feel free to refer back to the email.
10. Chase.com used or instructed chasesupport.com to send the email [agreement]
11. Please explain your choice

A.2 Codebook for Question

Table 5 and Table 6 show themes, codes and explanations for the questions that asking for interpretation of 'via'. This codebook coded the combined answer for the question "Please elaborate on what you think 'via' means" and " What information do you think Gmail is trying to communicate by showing 'via' for this email? "

Theme	Code	Explanation
Via indicates sender	Identifying the actual sender	Indicating or Identifying the actual sender (i.e., this is coming from chasesupport.com that's it)
Via indicates group association	From Chase's support division	Coming from chase' support division or mentioning the relationship between chase.com and chasesupport.com)
Via means through	Through a third party	Identifying that the email is sent through a third party or mailing list

Table 5: Codebook on questions asking for participants' interpretation of 'via' in the Support group

Table 7 and Table 8 below show themes, codes and explanations for the question "Why do you think Gmail has chosen to display 'via' to users for certain emails?" on the Support group and the Random group.

Theme	Code	Explanation
Via indicates sender	Identifying the actual sender	Indicating or identifying the actual sender or server that sent the emails
Via indicates group association	From Chase's support division	Coming from chase' support division (i.e. mentioning the relationship between chase.com and chasesupport.com)
Via encourages caution	Security or scam	Showing via to notify security reasons
	Forward address	Showing the forward address
Via means through	Through a third party	Showing that the email is sending through a third party or a mailing list
	Return-path domain	Identifying the domain as a return-path domain
Others	From a new contact	Identifying the email is from a new contact

Table 6: Codebook on questions asking for participants' interpretation of 'via' in the Random group

Theme	Code	Explanation
Gmail displays via for safety	Security	Showing for security reasons like phishing and legitimacy
Gmail displays via to inform you about the email	Outsourced	Specifically mentioning that the email is sent by a third party and not directly from Chase
	More info	Providing more information (e.g. where it comes from or on the sender) or showing to provide more transparency
Gmail always displays via	Showing reasons	Showing users why they are getting the email.
	Explaining via	Explaining that the email is sent on behalf of other sender
Unsure why Gmail displays via	Unsure	Don't know or unsure

Table 7: Codebook for the question: "Why do you think Gmail has chosen to display via" for the Support group

Theme	Code	Explanation
Gmail displays via for safety	Security	Showing for security reasons like phishing and legitimacy
Gmail displays via to inform you about the email	Outsourced	Specifically mentioning that the email is sent by a third party and not directly from Chase
	More info	Providing more information (e.g. where it comes from or on the sender) or showing to provide more transparency
	Authenticity	Adding the domain to show authenticity
Gmail always displays via	Explaining via	Explaining that the email is sent on behalf of other sender
Unsure why Gmail displays via	Unsure	Don't know or not sure

Table 8: Codebook for the question: Why do you think Gmail has chosen to display via for the Random group

Table 9 and Table 10 below show themes, codes and explanations for the reasons on the agreement questions "Chase.com used or instructed chasesupport.com to send the email" on the Support group and the Random group.

Theme	Code	Explanation
The via domain is suspicious	Scam	Email could be a scam or phishing email
	Suspicious	Email looks suspicious or the user feels concerned
That's just how it works	Sending service	The user identifies the random domain as a third party sending service or a bot
	Relevant domain	The domain name (chasesupport.com) seems related to chase
	Verified by Google	Google or Gmail verified that this email
	Via meaning	It conforms to the definition of 'via'
	Chase initiated	Chase initiated or approved the email correspondence
	Make sense	The explanation makes sense or conforms to users' mental model
	Unable to tell	The user cannot determine the relationship between chase and chasesupport. The user only knows that the actual sender is different than the purported sender
I'm not sure who did what	Unsure	Don't know or not sure

Table 9: Codebook for the reasons that participants agree or not agree that chase instructed the entity to send the email for the Support group

Theme	Code	Explanation
The via domain is suspicious	Scam	Email could be a scam or phishing email
	Suspicious	Email looks suspicious or the user feels concerned
That's just how it works	Sending service	The user identifies the random domain as a third party sending service or a bot
	The meaning of via	It conforms to the definition of 'via'
	Chase initiated	Chase initiated or approved the email correspondence
	Make sense	The explanation makes sense or conforms to users' mental model
I'm not sure who did what	Unsure	Don't know or unsure
Chase doesn't do this	Uncommon	It's an uncommon case or domain
	The other way around	The other way around (random instructed chase to send the email)

Table 10: Codebook for the reasons that participants agree or not agree that chase instructed the entity to send the email for the Random group