Assessing Suspicious Emails with Banner Warnings Among Blind and Low-Vision Users in Realistic Settings

Filipo Sharevski, *DePaul University;*
Aziz Zeidieh, *University of Illinois at Urbana-Champaign*

https://www.usenix.org/conference/usenixsecurity24/presentation/sharevski

This paper is included in the Proceedings of the 33rd USENIX Security Symposium.

August 14–16, 2024 • Philadelphia, PA, USA

978-1-939133-44-1

Open access to the Proceedings of the 33rd USENIX Security Symposium is sponsored by USENIX.

Assessing Suspicious Emails with Banner Warnings Among Blind and Low-Vision Users in Realistic Settings

Filipo Sharevski
DePaul University

Aziz Zeidieh
University of Illinois Urbana-Champaign

Abstract

Warning users about suspicious emails usually happens through visual interventions such as banners. Evidence from laboratory experiments shows that email banner warnings are unsuitable for blind and low-vision (BLV) users as they tend to miss or make no use of them. However, the laboratory settings preclude a full understanding of how BLV users would realistically behave around these banner warnings because the experiments don't use the individuals' *own* email addresses, devices, or emails of *their* choice. To address this limitation, we devised a study with n=21 BLV email users in *realistic settings*. Our findings indicate that this user population misses or makes no use of Gmail and Outlook banner warnings because these are implemented in a *narrow* sense, that is, (i) they allow access to the warning text without providing context relevant to the risk of associated email, and (ii) the formatting, together with the possible actions, is confusing as to how a user should deal with the email in question. To address these barriers, our participants proposed designs to accommodate the accessibility preferences and usability habits of individuals with visual disabilities according to their capabilities to engage with email banner warnings.

1 Introduction

An inclusive human-centered security design needs to accommodate the considerations of three dimensions: (1) *security*, (2) *usability* and (3) *accessibility* [25, 26]. Accessibility refers to the means that address discriminatory aspects related to equivalent user experience for people with disabilities, allowing them to equally perceive, understand, navigate, and interact with technologies, as well as contribute equally without barriers [39]. Relative to usability, accessibility demands that any human-centered security design does not disproportionally impact people with disabilities [3]. Relative to inclusivity, accessibility also demands that everyone is involved to the greatest extent possible, that is, factoring not just for their disabilities, but also their basic capabilities (e.g., access to the Internet, computer literacy, economic situation, etc.) when designing security technologies [5, 8]. While the security and usability dimensions are well established, accessibility has seldom been considered in a security context. Take phishing emails for example — there are plenty of studies considering the security and usability of warnings among sighted users [11, 24, 33, 36], but only a few consider the accessibility aspect for blind and low-vision (BLV) users [15, 18, 40].

Evidence shows that BLV users struggle with answering a large number of emails and are unable to distinguish warning pop-ups in their browsers while accessing email clients [15]. When accessing their email clients in browsers, BLV users are left to infer themselves the actions available as a result of a browser phishing warning (e.g., button to abandon a phishing link) based on ambiguous wording, ultimately exposing them to a higher risk of ignoring the warning altogether and falling pray to a phish [18]. Heeding phishing email warnings in clients like Gmail is also mired in difficulties. As the warnings' implementation hardly corresponds with the habits of using assistive technologies (e.g. using shortcut keys to skip over elements in the email) nor includes the necessary HTML tags (e.g., link, heading) for proper parsing, BLV users often miss the warnings altogether [40].

While this evidence is important in advancing the accessibility dimension of any human-centered security design, it is nonetheless limited to laboratory settings where BLV participants use email clients/browsers controlled by the researchers and are exposed to a set of either common phishing emails [18] or pre-selected combination of phishing and legitimate emails [40]. This setup precludes from understanding how BLV users behave under realistic conditions where they assess emails sent to their *own* email address, use different email providers for work (e.g. Outlook) and private communication (e.g. Gmail), or check their email through an app, a browser, or both. The email stimuli are of little relevance to BLV users personally because none of the emails are addressed to them, the content might be outdated, or the topic of the email is unfamiliar or inapplicable to them. Also, BLV users encounter spam emails in addition to phishing ones as

another type of potentially harmful communication, for which email providers also offer banner warnings in their clients.

An assessment of suspicious emails in a realistic scenario where all of these limitations are addressed at once is rather challenging but not prohibitive as past studies have observed the daily tasks BLV individuals perform when accessing their email clients [15]. To contextualize this observation when suspicious emails and the associated banner warnings are included, both for phishing and spam messages, we crafted a specific scenario in our study where BLV individuals used their *own* email correspondence sent to their *own* address and accessed through a client, device, and assistive technology of their *own* choice to answer the following research questions:

- **RQ1:** How a) accessible and b) usable are the phishing or spam banner warnings offered by email providers for BLV individuals in determining the legitimacy and the actions they take about suspicious emails sent to their own email addresses?
- **RQ2:** Do BLV individuals adhere to banner warnings about suspicious emails and to what extent do they continue to pay attention to a banner warning after encountering it multiple times?
- **RQ3:** How do BLV individuals assess suspicious — phishing or spam emails sent to their own email addresses (i.e. what cues do they use to determine the emails' legitimacy and what actions they perform)?
- **RQ4:** What recommendations do BLV individuals have for improvements in accessibility, design, and usability of banner warnings about suspicious emails sent to their own email addresses?

We obtained approval from our Institutional Review Board (IRB) to conduct a study with a sample of n=21 BLV individuals (original approval for 20 participants or above) who are regular email users (recruited through a snowballing technique). We set to perform our observations over Zoom where participants were asked to perform a few tasks with emails of their choice during which they verbalized their steps and provided their experiences/opinions. Respecting the privacy of their personal email correspondence in the main inbox, the first task was to review several emails of their choice in their spam/junk folder with their assistive technology of choice and verbalize to us only the subject line and sender. Using this verbalization, we ensured that these emails were not dangerous (e.g., resembled standard spam or phishing) both by checking for pretext and formatting patterns found in databases of known spam/phishing emails [23] and using our own assessment experience. If we were uncertain, we asked the participant to proceed to the next one (during the debriefing, we advised them that it is best to delete it). We then asked the participants to open the email, assess the legitimacy of the email (with a baseline established beforehand, as described in Appendix B), and share with us what the most likely action they would perform on it (without actually clicking any links or downloading any attachments).

To ensure that participants encountered both phishing and spam banner warnings (which are more prevalent in the spam/junk folders than the main inbox), we secretly sent a *false positive* phishing email to their own email address (the one they used to sign up for the study) — an email that providers deem as "phishing," assign a phishing banner warning, and automatically put it in the spam/junk folder — but in reality the email is legitimate. Once we completed the assessment tasks, we asked our participants about their accessibility and usability design recommendations for the banner warnings as well as their real-life experiences with detecting and acting on phishing and spam emails in general.

Our findings indicate that the email banner warnings in the two most prominent email clients, Gmail and Outlook, are implemented in a *narrow* sense. The *narrow accessibility* means that BLV users are only offered a mere verbalization of the warning text without conveying the reason why the banner is applied, the personal risks of engaging with the email, and the possible corrective actions relative to the automatic decision to filter the email and apply the warning. The *narrow usability* means that the warnings were technically noticeable but with a confusing textual formatting that is disconnected from the accessibility dimension in conveying what the banners are supposed to achieve security-wise for users whose needs, so far, have been excluded or overlooked.

This *narrow* sense of accessibility and usability compliance was perceived as a barrier to a *broader, usable accessibility* that is inclusive to the needs, habits, and experiences of BLV users with suspicious emails. To address the barriers of the narrow accessibility and usability of the real-world email banner warnings, our participants proposed designs that (i) mark suspicious emails as early warnings and allow BLV users to "*learn why*" a particular email was marked; (ii) allow configurable ping tones both "*before*" and "*after*" the text conveying the marking; (iii) a comprehensible text that conveys the risk level of marked emails; and (iv) different-than-rectangular shapes, in addition to stark contrast/and colorblind accessible colors for individuals who use screen magnifiers.

Contributions Our work has three contributions toward an inclusive human-centered security design:

- A novel methodology for conducting experiments in realistic conditions with security and safety-inducing interventions for emails with users whose needs have been, so far, excluded or overlooked in the current designs;
- Evidence of barriers to secure email engagement in realistic conditions due to *narrow* accessibility and *narrow* usability implementations (i.e., disconnected designs that don't account for the BLV users' overall email experience or basic email literacy);

- Recommendations for *broader, usable accessibility* redesign of email banner warnings (or other security warnings, in general) that consider the accessibility needs, usability habits, and basic capabilities of BLV users.

2 Background and Related Work

2.1 Usable Security Warnings

Usable security warnings are interventions in the user interfaces that are placed *in situ* with the objective of raising users' awareness of potential phishing attempts during their primary course of action [11]. They usually come in a couple of variants, depending on whether they require user interaction or not. The first variant is *interactive* warnings that either force the attention of a user to a suspicious element like a link or attachment (usually within an email client) [24, 36] or offer options for users to "go back to safety" or "proceed to a website" (usually in a browser). Often, the interactive warnings come with additional informative messages that communicate the threats ahead in an effort to proactively raise awareness for future cases of phishing exposure. The second variant is *passive* warnings that attempt to indicate to a user an imminent phishing warning without interrupting their primary course of action either by highlighting a domain or chaining the color of a browser's address bar.

The evidence from user evaluations of both types of warnings suggests that users often ignore the passive warnings, but they tend to adhere to the interactive warnings [11], provided the wording is comprehensible and incorporates a design that prevents habituation (i.e. attenuation of a user interaction response with multiple exposures to a same warning) [33]. The adherence and phishing safety, however, does not come without a cost – usually, the forced attention is distracting, time-consuming, and tedious [38], especially with a high number of emails a user receives a day, and the element of fear increases with the repeated exposure to decisions to abandon a suspicious website [27]. There is also a difference in effectiveness whether the warning "friction" happens within an email client as a banner (the usual vector for delivery of phishing attacks [34]) or in a browser as a splash screen, with the later implementation being better preventing participants from reaching phishing websites [24].

2.2 Access to Usable Security Warnings

The evaluation of usable security warnings usually is done with *sighted* users or users who don't need an assistive technology to access the warnings. But the technologies that these warnings attempt to prevent from security failures necessitate a material access with technical capabilities and know-how, something that is not available or intrinsic to all users [19]. A digital divide thus exists between users [28] that disproportionately renders many of those without access and digital literacy to greater security and privacy risks [3, 4]. This "digital exclusion" often brings on, or exacerbates, the vulnerabilities of marginalized and underserved user groups [26].

Vulnerability in this context arises from the lack of access to resources or barriers to using a security or privacy technology in the way for which it is designed. These are "human vulnerabilities" and differ from the usual "cybersecurity vulnerabilities" (i.e., weaknesses in the technologies that result in security protections' failure [22]), though the notorious statement "humans are the weakest link" treats all users as the cybersecurity vulnerabilities. In response to this (mis)treatment, a user-centered security approach has been adopted [27, 41] and, accounting for the human vulnerabilities, an inclusive security approach has recently gained traction [37]. The latter approach brings the attention of evaluation, involvement, and empowerment of users with human vulnerabilities with security technologies, such as authentication [12], usable security warnings [18, 40], and misinformation [30].

The inclusion pertains to users with physical vulnerabilities (such as blindness, motor impairments, deafness, hearing loss, etc.), cognitive (such as memory loss, mental health conditions, etc.), financial vulnerabilities (such as unemployment, homelessness, etc.), and emotional (such as fear appeal, upsetting security and privacy experiences, etc.) [26]. We, in our work, were particularly interested in studying the exclusion of BLV users when designing usable security warnings that preclude access to people with visual disabilities. From a visual accessibility perspective, it appears that the design is driven towards "visual" frictions intended to enhance the security behavior of sighted users [9].

The objective of forcing attention, for example, is achieved by displaying (i) interactive warnings that appear when a user hovers over a phishing link; (ii) a banner between the email header and body; or (iii) a splash screen with an informative message that offers users the option to click back to safety or proceed to a website. In the first case, a sighted user has no problem hovering over a hyperlink in an email, but that is not the case for a BLV individual who relies on a screen reader to verbalize the contents of an email. In the second and third cases, the banner is often implemented in red color and contains signage like stop or exclamation mark signs to highlight the danger of phishing ahead. Red as a color holds little significance to a (color) blind person who is also unable to fully capture the warning signage when accessing the email in a text-to-speech fashion.

The promising utility of the interactive warnings comes short when it comes to accessibility as evidence further shows that the options for "going back to safety" or "proceeding to a website" are incongruent between different browser implementations and often confuse BLV users about what is the outcome of the interactive element in the warning [18]. The email banners warning equally suffer from inconsistency when it comes to standard and HTML presentations and though verbalized by screen readers, they make a very weak case of

adherence [40]. It's not that the accessibility is insufficiently or partially addressed when it comes to communicating the message of the warning — the design entirely omits to incorporate the email and browser habits of BLV users too.

Usability-wise, BLV users tend to escape from non-usable or inaccessible content by tabbing or scrolling down [35]. This user population usually "scans" a website with their ears by listening at a high speed and rapidly exploring the page by jumping directly to headings and links through heading lists or link lists provided by the screen reader [32]. Blind or low vision users often remember the "metadata" that needs to be skipped in order to reach their desired content on a website of their interest [35]. Also, rich, well-formulated textual content seems more credible to blind or low vision users even if the visual appeal seems suspicious to visually able users [1].

2.3 Usable Security Warnings and BLV Users

An evaluation of warnings about dangerous emails, so far, has mainly been performed relative to their accessibility [18, 40]. In [18], BLV users were shown an example of a phishing email sourced from a repository of common phishing emails [6] and informed that if a user clicks on the links inside the email, a standard browser warning will be displayed with a description of the phishing risk ahead and the options to abandon the link (e.g., "Go Back" button), proceed (e.g., "advanced" link), or learn more. Then, the users proceed to evaluate this warning with their own assistive technology by opening a researcher-controlled link in a browser of their choice.

Though this approach is certainly better than a laboratory set security warning evaluation, it has a couple of limitations. First, participants are shown an email that neither is addressed to them (i.e., to their *own* email address) nor the content is of significant relevance to them. While the emails, coming from Cornell University's repository [6], employ the known elements for luring victims (e.g., authority, scarcity, etc.), they usually use a pretext relevant to academic environments (e.g., a campus job offering, university account expiring, etc.), which might neither be of interest to participants nor would they be targeted with such emails. Second, these emails, if not filtered by default, usually invoke banner warnings in the email client of choice of the participants that precede the browser warnings. These banner warnings together with the email itself, even with a variable degree of accessibility, might dissuade a BLV user under realistic conditions from clicking the link and accessing the browser altogether.

An evaluation of email client warnings with BLV users was performed in [40]. Here, the emails were also selected from a phishing email repository and were spoofing three popular email senders including Google, Apple, and Hulu. The emails were placed in a Gmail account controlled by the researchers and opened in a Chrome browser, that was running in a virtual machine with a preloaded assistive technology of the choice of each participant. Participants accessed this setup remotely through Zoom (screen sharing and remote control function) and played the role of an assistant to help their manager process and review emails. The emails triggered an icon warning that displays a red question mark in the sender's avatar icon of the email and, on hovering over, pops up a short message: "Gmail couldn't verify [sender address] actually sent this message (and not a spammer)."

While this setup in [40] was created to test a baseline behavior that later would allow comparison with their implementation of inclusive email security indicators (a very commendable and welcomed contribution), it nonetheless has several limitations towards a realistic evaluation of email security warnings. As previously noted, the selection of the emails might not be relevant for the participants as they are not addressed to them personally. Next, many BLV people use Outlook or other email clients in addition to Gmail, and they might check their emails through their phone's applications in addition to a browser. A role-play scenario in which a BLV individual plays a manager's assistant might also be something that many of the participants are unfamiliar with or it might not represent their typical daily engagement with emails. Importantly, the icon tested in the baseline scenario differs from the email banner warnings that are usually triggered for both phishing (example shown in Figure 1) and spam (example shown in Figure 2, and tested neither in [18] nor in [40]) and contain much more a verbose text of the warning compared to the icon's alt text (particularly an important feature for users that rely on aural email access).

3 Study Methodology

Situating the study in realistic settings entailed a specific protocol that ensured our participants were not exposed to greater than minimal risk but also using mild deception to avoid priming them about exposure to real-world phishing and spam emails (the entire debriefing about this deception is given in Appendix D). Instead of exposing participants to phishing emails that weren't addressed to them or using a laboratory setup and a role play scenario for accessing these emails, we invited our participants to join a 45-minute, audio-only, recorded Zoom interview session where they can use their *own* device, their *own* email client, and their *own* assistive technology as they go about in their everyday lives with their *own* email address. We avoided asking them to go through their primary inbox for several reasons, namely because: (i) this might have constituted a greater than minimal risk of privacy invasion for which we did not have available safeguards to prevent; (ii) the probability that a participant will encounter a phishing or spam email with warnings during a brief sorting task of their inbox during the interview time is very low; and (iii) major email providers might display banner warnings, but automatically move untrustworthy emails to the spam/junk folder by default [14, 20].

To address this challenge, we decided to direct our partici-

pants to access several emails within their spam/junk folder. This would have constituted a priming to phishing/spam if we had done so at the beginning of the study, therefore, we asked them a couple of questions prior to starting efforts to frame the study in a broader suspicious email context. The first question was about how often our participants receive suspicious emails in their inbox folder (not caught by a spam/junk filter) and the second was about how often they encounter legitimate emails in their spam/junk folders. These answers allowed us to provide the opportunity to frame the subsequent request to sort several emails from the spam/junk folders as a task towards assessment of suspicious emails that are not expected to be phishing or spam by default just because they have been filtered as such by the email providers. Once in the folder, we allowed our participants to select any emails they wanted, but we had to ensure a higher probability that each participant would encounter phishing, spam, or both emails — with their associated client banner warnings attached by providers shown in Figure 1 and 2 (Gmail) and Figure 3 (Outlook), respectively — during the Zoom interview session.

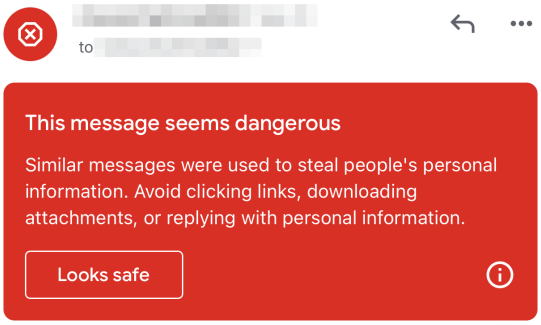

Figure 1: Gmail phishing warning banner

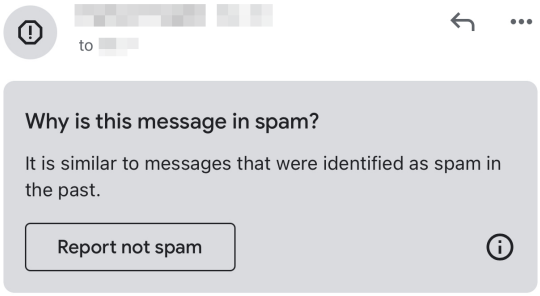

Figure 2: Gmail spam warning banner

While a prior check of multiple researcher-controlled email addresses revealed to us that encountering a spam email with the associated banner would not be a problem, that was not the case for encountering a phishing email in the participants' spam/junk folders. To ensure that this would happen for the

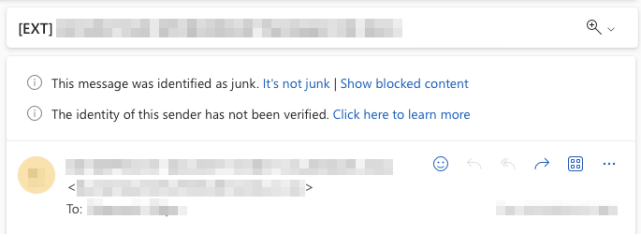

Figure 3: Outlook warning banner in the junk folder (note: Outlook uses the same for any untrustworthy email without distinction of the email type between phishing and spam)

purpose of the study, we have prepared a *false positive* phishing email that we knew the filters would assign a phishing or banner warning and move to the spam/junk folder. We decided to use a false positive phishing email, that is, an email which was *not* phishing but was classified as such anyhow in order to avoid exposing our participants to greater than a minimal risk with other, *true positive* phishing emails. This means the email could not cause any harm to BLV users as it did not contain any malicious attachments or links involving an unduly phishing risk (all links lead to legitimate Amazon Web Services (AWS) pages). As a *false positive* phishing email in our study we initiated an AWS account creation verification email, shown in Figure 4, directed on the day of participation to the participants' emails they used to sign up for the study (we didn't want to tip them off that something might be amiss if we sent the email right before their Zoom session).

We chose this email as we encountered it in our own spam/junk inboxes after we did a personal AWS initiation. We refer to it as "mild deception" as our participants did not know we were the ones that instigated the sending of this email to their addresses. We were fairly confident that the participants would already have spam and we were prepared to use the spam instance if needed, but we were less sure about them having phishing emails sitting in their spam/junk folders so we could also test the phishing warning banner variant. The AWS email was necessary to invoke a realistic scenario where our participants access a phishing banner warning assigned to an email because such an occurrence might not happen frequently enough to be reasonably observed as part of the participants' typical email engagement during the study. To ensure the AWS email will work for all participants (Gmail will assign the warning banner and put it in the spam/junk inbox), we created 10 different email accounts, sent the AWS email to them, and found that to be the case in all of them.

We were aware that the classification of untrustworthy emails was predicated on the individual's email correspondence and behavior, and we expected that we might encounter a case where the AWS email might not end up in participants' spam/junk folders. For those cases, we decided to proceed only with what they had in their spam folder as emails addressed to them without going to the main inbox or attempting

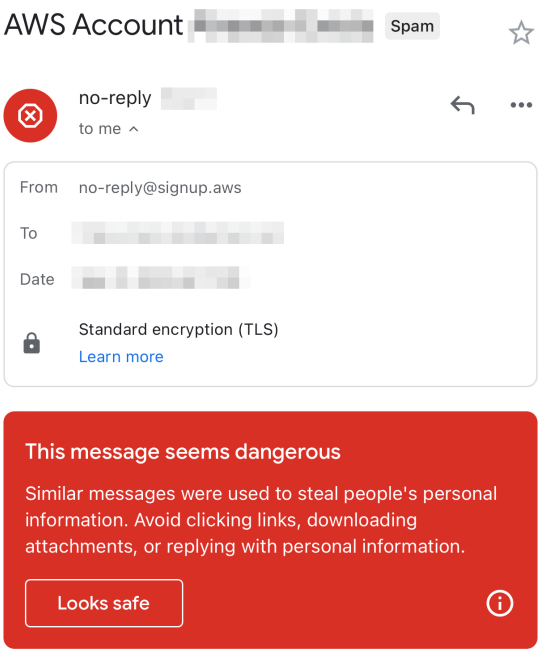

Figure 4: The *false positive* AWS sign-up email

to ask them to perform additional steps. The study has been approved by our IRB and we used an extensive debriefing (see Appendix D) in which we pointed out our methodology, discussed any surprises or events during the interviews where we felt participants might be in the line of phishing/spam harm in the future, and pointed participants to general email phishing resources if they wished to raise awareness on the subject. We believe that our overall methodological approach is appropriate because we strike a good balance between the acceptability/manageability of our participants to participate under realistic conditions while we grasp their real-time experiences with their personal email correspondence.

3.1 Participant Recruitment

For our recruitment, we followed Gerber's advice when doing usability and accessibility studies with blind and low vision people [13]. Accordingly, we recruited participants who use a screen reader, screen magnifier, or both, that regularly use web/email clients, and could obtain online information aurally without problems. They had to be individuals 18 years of age or older who had internet access on their own device, client, and browser and were English-speaking and literate. As one of the researchers is a legally blind individual, we recruited potential participants through snowballing where we partially sampled personal acquaintances and partially a pool of BLV participants that was recommended by one of our acquaintances. We used a formal email approved by our IRB (see Appendix A) to approach each of the potential participants. We arranged audio-only recorded Zoom interviews with interested respondents on a rolling basis, requesting that each participant has access to their preferred way of using emails and their assistive technology. The interviews lasted on average 45 minutes, and we compensated each participant with a $25 Amazon eGift Card ($525 in total).

3.2 Trust and Ethical Considerations

As this was a study in realistic settings and concerning the participants' own email addresses, it was important for us to establish trust and assurances about the goals of the study and the safeguard protections we had in place. Following the suggestions for considering the ethical aspects when doing research with blind or low vision individuals from prior work [16] we first obtained verbal consent both before we started the audio recording and afterward (to have evidence in our transcripts, but also to avoid creating a recording in case a potential participant does not consent, in which we would have thank them and closed the Zoom session). Next, we communicated that the goal of our study was to capture the "richness" of their experiences in realistic settings rather than the commonality of these experiences in laboratory settings. This is the reason, hence, why we are asking them to work with their own email address, client, and access technology. Here, we offered the option for them to choose not to participate and to choose which address they will use — as many of them have multiple email addresses (e.g., work, private, subscriptions). We noted they could select any email they wanted to communicate with us from their spam/junk inbox and they were free to stop and abandon any question at any point in time if they felt doing so. Prior to doing any of the tasks and the interview, we told them that they could ask us to stop the interview, stop the recording, or remove any answers or readings at any point in time.

Only after we received the participants' explicit permission that they were okay with proceeding with the study and accessing the spam/junk folder of their own email address, we commenced the audio-only recorded Zoom session and proceeded to complete the tasks and the interviews. Following the recommendations for doing usable security research with BLV users [29], we verbally notified each participant when we started each audio recording and we explicitly told them that they could take as much time as they needed to answer any question. We allowed them to verbalize the process, give comments, complaints, suggestions, and verbalize any other experience that was not necessarily with emails but other technologies (e.g., social media, passwords, software) in order to allow for them to fully express the natural behavior "surrounding" their daily interactions with email.

We also pointed out to our participants that they could act on the emails from the study as they ultimately wished (e.g., delete, move to inbox, report, etc.). After the we collected participants' answers, we verbosely debriefed the participants about the mild deception we used and that we were the ones

that initiated a *false positive* phishing email to their spam/junk email (if they received one or chosen to verbalize one during the study). We pointed out participants to general email phishing resources if they wished to further raise or check their awareness [7]. To ensure we obtained a correct understanding of their experiences and recommendations, we reviewed the main points we recorded during the interview and clarified any misunderstandings we might have. We also sent a draft of our paper to our participants for feedback.

We employed lengthy explanations to ensure our participants that we were not involved with the filtering nor with the formatting, verbalization, and accessibility implementations of the email banner warnings they saw during the study or in the past. We were also careful not to appear in favor nor support of particular types of warnings in order to maintain full researcher impartiality. We communicated that our ultimate goal is to create meaningful *inclusion* for BLV individuals when these banner warnings are present in their *own* email clients of choice. We pointed out that, this goal, however, doesn't prevent from misusing our findings or misinterpreting them in making compromises for accessibility or removing such support altogether, given the trend of big tech companies to do just that for cost saving reasons [17].

3.3 Data Collection

Initially, the interview transcripts from our Zoom sessions were not anonymized, but we removed any names and references to individual participants and deleted the audio recordings altogether. The transcripts, assigned only with a participant number in the order of participation, were stored on a secure server that only the researchers had access to. Each interview was done with open-ended questions, listed in the interview script (see Appendix B). Due to the nature of the study, not all the participants used a similar set of emails, and in some cases, the email providers either entirely filtered out the *false positive* phishing email (shown in Figure 4) or moved it to the main inbox without assigning any banner warnings.

We concluded our recruitment with a sample of 21 participants as we reached thematic/data saturation (i.e., we collected data up to the point where there were fewer surprises in the responses to the research questions and no more emergent patterns). As part of the debriefing process, participants were offered the option to withdraw from the study after finding out about the mild deception (i.e., the *false positive* phishing email) or no later than 30 days period after the data collection concluded (none of the participants exercised this option). The demographics, email setup, and the sample's visual profile, per the suggestion in [13], are all given in Table 1.

Gender				
Female 11	Male 9	**Non-Binary** 1		
Racial/Ethnic Self Identification				
White 14	Latinx 3	Asian 1	Black 1	Other 2
Age				
[18-29] 2	[30-39] 7	[40-49] 7	[50-59] 5	
Education				
High-school 1	College 18	Post-Graduate 2		
Visual Self Identification				
Totally Blind 4	Blind, Perception of Light 10	Low Vision 7		
Device				
iPhone 12	Laptop/Windows PC 9			
Provider				
Gmail 19	Outlook 2			
Client				
App 16	Web 5			
Assistive Technology				
Screen Reader 18	Magnifier 3			

Table 1: Demographic Distribution

3.4 Data Analysis

Since we had to work with a degree of arbitrary selection of emails in our study, we asked our participants to provide lengthy responses to our questions and asked for further clarifications. With the collected data, we performed an inductive coding approach to identify frequent, dominant, or significant aspects of their answers. As suggested in [2], we first familiarized ourselves with the data as we had to manually revise each transcript to remove personally identifiable information. Next, we completed a round of open coding for arbitrarily selected two interviews to capture the participants' decision-making process around the tasks they performed in the study. Then we discussed the individual coding schemes and converged on an agreed codebook. We used this codebook to independently code the remaining interviews, reaching an *Inter-Rater Reliability* (IRR) of $k = 0.9$ (Cohen's kappa), which we deemed acceptable. The themes we identified respective to the research questions, were then discussed, interpreted, and example quotations were selected to represent each of the findings [10].

The coding was informed by a shared accessibility experience relative to what it means to pay attention when using a screen reader and the effort needed when accessing emails using assistive technology. This helped us resolve mismatches that usually arose from differences in familiarity with a particular screen reader or other assistive technology. The codebook (see Appendix C), thus, captured four main aspects: (i) *email banner warning accessibility and usability* i.e., codes related

to the accessibility and the phishing-avoidance utility of the email banner warnings; (ii) *email banner warning adherence and attention* i.e., codes related to the adherence and attention paid to the email banner warnings; (iii) *email assessment* i.e., codes pertaining to cues, criteria or rules of thumb used to determine the legitimacy of personal emails; and (iv) *accessibility and usability redesign* i.e., codes describing the participants' recommendations for accessibility and usability redesign of the email banner warnings towards a safer use of emails by blind or low vision users.

4 Results

As this was a study in naturalistic settings, participants had the opportunity to choose and select which and how many emails they would like to access and assess in their spam/junk folder. As previously mentioned, we initiated the *false positive* phishing email to each of the participants on the day of the Zoom session, but we were aware that it might not end up in their junk/spam inbox because of personalized filtering or because the participant might have deleted it before. Table 2 gives the breakdown of the number of emails, types, and banner warnings encountered by each participant.

P#	Gmail Spam	Gmail Phishing	Outlook Junk
P1	1	1	n/a
P2	1	1	n/a
P3	1	1	n/a
P4	2	1	n/a
P5	2	0	n/a
P6	1	1	n/a
P7	2	1	n/a
P8	2	1	n/a
P9	2	1	n/a
P10	n/a	n/a	2
P11	2	1	n/a
P12	2	1	n/a
P13	n/a	n/a	2
P14	2	1	n/a
P15	2	1	n/a
P16	2	0	n/a
P17	1	2	n/a
P18	3	0	n/a
P19	3	0	n/a
P20	2	1	n/a
P21	3	0	n/a

Table 2: Personal Emails Used in the Study per Participant

4.1 Study Framing

Equally, as their sighted counterparts, BLV users do receive suspicious emails in their inbox folder (not caught by a spam/junk filter) – some of them receive a few a month, some of them rarely, and some of them all the time. These emails revolved either around "*password reset*" or a work-from-home opportunity where the email receiver is free to "*make their own hours and make $95 an hour*" (**P10**). The senders of these emails ranged from businesses, banks, insurance companies, and service providers to simply individuals from the perspective of "cognitive triggers" of phishing compliance, blind and low vision users are targeted with the *scarcity* and *reciprocity* pretexts, reinforced with *authority* as the alleged email initiator. In an attempt to trick spam/junk filters, phishing attacks sometimes contain only attachments, one of which is an image of what should have been the body of the email. This phishing tactic, for BLV users proves hardly efficient because, as **P21** put it, their "*screen reader was reading neither the sender nor the attachment content*, " which was a cue that the overall email was "*super phishy*." Interestingly, this might be a phishing case where BLV users, have an advantage, in the form of their screen reader, compared to sighted users.

A similar outcome was observed relative to the encounters of legitimate emails in BLV users' spam/junk folders. This happens regularly, and our participants stated that these are often important emails that they found following a "*check your spam folder*" reminder email in their main inbox, as participant **P11** noted. Without this reminder, participants complained that there is little to be done because "*marking these emails as 'not spam/junk'*" (**P11**) or "*putting these email addresses in the address book*" (**P4**) doesn't prevent them missing important correspondence. The problem is exacerbated when it comes to 2-Factor Authentication (2FA) because, according to participant **P12**, "*it delays their login, causing them to make additional phone calls*." Considering security protections from a broader usability perspective, this finding reveals that this machine-assisted design of filtering suspicious emails negatively impacts people with visual disabilities [3] as it induces a burden in using a second layer authentication as an enhanced protection.

4.2 Accessibility and Usability Evaluation

To address the gap in knowledge of how BLV users engage with email banner warnings under realistic conditions [18, 40], our first research question looked at the accessibility and usability of these warnings when applied to the participants' own emails by their preferred email provider.

4.2.1 Gmail *Phishing* Banner Warning

Out of the 19 participants that used Gmail, either on their phone or laptop, 14 received the *false positive* AWS email shown in Figure 4. One participant also received another phishing email in addition to the *false positive* AWS email (**P17**). The first thing we discovered was a discrepancy between the visual formatting in the friction and the formatting

delivered through a screen reader. Several participants mentioned that their screen reader uttered the word "*phishing*" before it read the sender and the subject line of the email, a feature not available for other assistive technologies such as screen magnifiers or otherwise in general (nor is the word phishing is found anywhere in the banner text). Usability-wise, participants didn't mention that this "hidden" accessibility feature of Gmail was much of a help because "*their attention is usually focused on the sender and the subject line*" (*P9*) and the "*phishing*" utterance was not an element they expect or are used to when reading emails.

Most of the participants using screen readers and all using magnifiers were able to access the Gmail phishing banner warning in a *narrow* sense, that is, they were able to hear the text of the warning or see the banner through magnification. But they also commented that "*it's not obvious [the banner] is a separate interactive element and there is no way to easily find it if you quickly went past it*" (**P9**). More importantly, the Gmail phishing banner did not address the accessibility in a *broader* sense, i.e., considering the usability and security aspects of these warnings – many participants missed or ignored them altogether when going through their emails. We had to explicitly point them back to the banners so they could go over again and access them again. This interactive drawback was also emphasized as a barrier precluding engagement with protections that should not disproportionately discriminate between sighted and BLV users.

Many of them, like participant **P3**, noted that they use their screen reader "*set on a much faster rate than a generally spoken text*" and, as such, they have no way of telling whether the banner warning text "*is part of the email's header or not*." Participants also commented on the lack of a literacy baseline that joins the accessibility and usability of the phishing banner warnings. For example, participant **P7** noted that the "*apprehension of the text is confusing*" because it doesn't say "*who's stealing my account or who's trying to steal my personal information*" or if "*somehow related to my other similar emails from Amazon*." Along these lines, Interestingly, here, participant **P3** even commented that the button "*looks safe*" means little to a user that *technically cannot use their sight to actually look*."

The nonintuitive accessibility was equally met with the nonintuitive usability of Gmail's phishing banner warning. Participants did mention that the additional text they heard '*give them like a little quick second to actually think what to do with an email*" (**P1**) or ""*delete the email straight away*" (**P2**), but they were aware that this *narrow* sense of usability is not sufficient to always induce a safe behavior. Participants, aware that "*Gmail doesn't always do a good job of filtering*" (**P3**), complained that they have to "*scroll through and come back a few times before deciding what to do with the email*." The lack of phishing warning literacy or an explicit agreement on what the banners are intended to achieve security-wise for users was also brought up as a barrier for a *broader, usable accessibility*. For example, the emphasis on "*danger*" without an explanation of what are the consequences of having your personal information stolen was perceived as an effort to elicit an "*emotional response*" (**P6**) without "*really communicating the severity of how harmful an email is*" (**P20**). This usability drawback, in turn, did not necessarily "*dissuade [them] from looking in each of those emails in detail*" (**P16**)

4.2.2 Gmail *Spam* Banner Warning

Participants, depending on the nature of the email, either accessed one, two, or three two spam emails with the warning banner from Figure 2. Here too, some of the participants missed the spam banner warnings and commented on the *narrow* sense of accessibility that precludes full user engagement with them. For example, **P18** noted that "*they are accessible as it's just text, but I kind of passed over them really quickly without paying much attention*." The absence of baseline literacy of how these help users was also pointed out as problematic. Here, participant **P21** commented that the "*warning is accessible, but the button 'report not spam' is confusing as it doesn't say to whom this report goes and whether the reporting will automatically unblock all future emails from this sender to never go to spam*." The *narrow* sense of accessibility was also present here, as participants "*often have to poke around to find what really is the beginning of the email message — the warning or the subject line*" (**P15**).

The nonintuitive usability of Gmail's spam banner warning was also brought up as an issue. The evolution of the spam warnings design – especially the wording that is intended to induce a safe behavior – wasn't helpful either, because "*previously [it] was saying to be careful with an email and now [it] just says why the email is in the spam*." (**P20**). Here, participant **P12** pointed to the confusing wording of the spam warning relative to the "similarity" reference inside: "*I understand [what it supposed to tell me], but it didn't make sense, because it's not something I've experienced before with my 'own' similar messages*." This usability drawback, similar to the Gmail phishing banner warning, did not "*dissuade*" the users to assess the email body nonetheless because "*the email might be something important to [them]*" (**P16**). This course or (in)secure action, in effect, was a result of the formatting of the warning's text itself (and the overall design) as it made the participants "*read the email anyhow to know whether to report it as spam or not*" (**P6**).

4.2.3 Outlook *Junk* Banner Warning

The participants that preferred the Outlook client did receive the *false positive* email from Figure 4 but the warning banner shown in Figure 3 was the same for it and for the other junk email that was in their folder. Overall, Outlook didn't fare any better than Gmail in enabling a *broader* accessibility for blind or low vision users that contextualizes the usability and

security of the "junk" email warnings towards a truly secure behavior under realistic conditions. Participant **P13** considered the Outlook banner warning in the context of what they habitually use in their main inbox, stating they "*have easily missed it or went pass through it because it appears after the email subject*" (unlike the Gmail's banners). The wording of the warning – as the main element of communication with users strictly dependent on aural guidance – showed to be a problem here too because it was "*too cumbersome and confusing in the first sentence and the follow-up link*" (**P13**) leaving users to wonder whether the email "*was or was not junk*." While the participants were able to navigate to the warning itself, this course of action didn't help them much because it did not preclude them from wondering *who* is the one that "*identified [an email they received] as junk*" – the provider, the email owner, or someone else (**P10**).

4.3 Adherence and Attention Evaluation

With the second research question we wanted to learn whether BLV users (i) adhere to banner warnings like this in general, and (ii) continue to pay attention to them after encountering them multiple times under realistic conditions. These aspects have been evaluated with sighted users in the past [11, 27], but have not been addressed in the previous work with BLV individuals [18, 40]. Participants confirmed they do adhere to the warnings mostly because they "*trust Gmail to filter the spam and phishing out for them*" (**P5**), even "*if there are instances when important correspondence has been sent to their spam inbox*" (**P21**).

But the trust-based adherence alone was not something that participants relied on when engaging with suspicious emails in their daily lives and instead relied on their "*own judgment*" (**P2**). Participants were wary that the implementation of the warnings wasn't catering to their accessibility needs (e.g., "*mentioning a 'danger' without clear danger ahead is confusing*"(**P4**)) or to their usability needs (e.g., the "'*report not spam' feature really doesn't solve the problem with one click and is ineffective of avoiding putting important emails in the spam folder time and again*" (**P4**)). The adherence exceptions that sighted users exhibited with security warnings [38] were also mentioned by the BLV participants in our study. Participants complained about the "*tediousness of getting legitimate emails from spam back to inbox*" (**P15**), which for a blind user is "*time-consuming*" (**P12**), and goes counter the intention of the warnings to help users get an "*intuitive sense*" of the phishing/spam risk ahead.

When it came to attention, we wanted to know whether the BLV users' interaction response attenuates with multiple exposures to the same warning. We couldn't use the same evaluation strategies for sighted users [33] so we asked the participants who accessed more than one spam (Gmail) or junk (Outlook) email about how attention-grabbing was the experience when they heard the warnings multiple times. Granted, we couldn't explicitly measure any "attenuation" effect, but we nonetheless uncovered the tendency of the BLV participants to avoid "*paying much attention to the warnings after the first one*" (**P21**). Their focus, instead, was "*more so to the subject and the sender to get the feeling if the emails are not actually spam*" (**P21**). Instead of heeding the warning each and every time, participants attention was driven towards finding a way to "*report an email to make sure that they don't get these emails again (automatically deleted)*" (**P19**) so they didn't have to deal with warnings altogether.

The narrow accessibility of warnings also played a role in the lack of attention to repeated warnings with our participants. The aural aspect of the warnings, unsupported by the color and signage for sighted users, makes them less attentive. The repetition of the warning text when using fast speed settings on a screen reader, according to participant **P18**, is of little to no help because "*keywords like 'spam' or 'phishing' are eclipsed by the convoluted text of the warning itself*" so the attention is shifted to the next email element instead. The narrow usability of the warnings, too, turned some of the participants to "*default being suspicious of certain emails in their junk folder*" and pay no attention to the warnings at all, regardless of any repeated exposure.

4.4 Suspicious Email Assessment Strategies

One of the most dominant cues of an email's untrustworthiness that our participants relied on was the aural inconsistency and grammatical errors resulting from typos, punctuation, and unexpected symbols. Our BLV participants pointed to their aural ability to spot misspellings, grammatical errors, typos, and inconsistencies in the use of symbols. The cues that the BLV participants in our study relied on – regardless of the presence or absence of warning banners – included "*random letters and numbers and 'weird' addresses*" (**P17**), "*grammatical errors*" (**P19**), and "*spelling errors, omitted letters, or extra letters*" (**P18**). Participants, like **P5**, noted that the aural cues of inconsistency make them distinguish between spam and other emails because "*spam seems to have a lot more punctuation or symbols instead of letters*."

The lack of warning context for BLV users, leaves them to rely on expected "*email narratives*" as a clue for potential phishing, spam, or junk emails. For example, participant **P16** mentioned that past experience accessing legitimate emails using assistive technologies helped them "*tell [when] the narrative sit right with [them]*" or not. The improbability of a request, for example "*an expired account, change of password, or attempted delivery*," was "*screaming of phishing*" – in the words of participant **P18** – because they knew when to change a password or what they have ordered. Here, the text of the warnings was of little help because the design is "*too general and doesn't mention examples such as passwords when it speaks of 'personal' information*" (**P9**). The banner warnings were irrelevant elements in detecting spam or phish-

ing when a request was inapplicable to the conditions of the participants. For example, participants **P11** shared that "*you know it's phishing or spam email when they try to sell you a car insurance but they don't know you can't drive a car*."

To the point of absent baseline literacy that joins the accessibility and the usability of the banner warnings, more than 70% of our participants mentioned they never received any formal training in phishing detection, including heeding warnings with accessible technologies. Those that did, mentioned a standard form of training like "*training modules, quizzes, and little scenarios asking 'should you open this email or not'*" (**P11**) that were not specifically considerate of the accessibility dimension of engaging with suspicious emails. This omission negatively affected our participants in reality as six of them disclosed they had been phishing or spam victims in the past. Participant **P8** said they fell for a phish and wondered whether a *broadly* accessible warning would have precluded them avoid an "*email requesting a change of the expiration date of a payment method*" because the phishing warnings shown in Figure 1 or Figure 3 don't explicitly refer to these details as 'personal information' nor to actions past 'clicking a link' such as changing expiration dates.

4.5 Design Input and Recommendations

Suspecting that the email banner warnings, in realistic conditions, might not particularly cater for aural reception and usability through assistive technologies, we asked our participants to share their preferences in how they would like to be warned about potential phishing spam or junk email. Past research with BLV users has successfully tested accessible warnings for screen readers that contain a bell sound, followed by a short introduction speech "warning! for details, press a shortcut key." [40]. The shortcut key, in turn, presented an overlay message that conveyed an overlay text of a general Gmail warning that urges users to be careful with the email message. Testing security warnings with suspicious content, though on social media instead of emails, blind and low vision users proposed several design options for placement of these warnings in order for them to be noticeable, comprehensible, and usable [30]. In agreement with the overlay warning approach from [40], these options included: (i) covers or banners that incorporate text that explains why the cover/banner is there in the first place and how the cover/banner content relates to a particular user's email/content; (ii) an audio signal or vibration that precedes the banner/content; and (iii) stark contrast, bold/large font, and standout colors that differ from the client/platform aesthetics so users dependent on magnifiers (or colorblind users) could notice the warning itself.

As we wanted to capitalize on the participants' immediate realistic experience with email banner warnings, in the four research questions we asked them to provide design input and recommendations that aim to address the interactive drawbacks they experienced through the tasks and reflections part of the previous research questions. Participants unanimously agreed that the current implementation of the real-life email banner warnings is accessible in *narrow* sense, i.e., it allows for a simple text-to-speech translation by a screen reader or presentation through a magnifier tool. Gmail's "*phishing*" utterance was seen as a potential opportunity to expand the accessibility in a *broader* sense, i.e., one that considers the usability and security aspects of the entire interactive warning experience for BLV people. For example, participants proposed replacing the "*phishing*" utterance with the wording "*Gmail marked this email as a phishing/spam; tab to see why*" (**P5**), and then on the tab action, a full warning "*explaining the features and the decision making of the particular email message that got it marked*."

The emphasis on explicitly "*tab to see why*" (**P15**), in participant **P9** view (and similarly to the shortcut key feature proposed in [40]), was to provide a baseline literacy that joins the accessibility and the usability of the phishing banner warnings. Participants reasoned a "*full transparency on filtering*" will help them better shape their engagement with emails in cases where Gmail/Outlook mistakenly apply (or omit) banners and puts the messages in a wrong inbox. Here, participant **P7** reasoned that if Gmail tells them that a "*mangled domain such as 'arnazon.com' instead of 'amazon.com' have been detected in this email*" then they would look for this cue for future emails, such as was the AWS *false positive* email we sent. sighted users, studies show, are able to benefit from such warning information [21]. A variant of the proposed warning alert that included a risk categorization level – akin to the defense readiness condition scale DEFCON – was proposed by participant **P7**. They reasoned that a "*numeric or categorization system*" would help BLV users "*quickly assess*" the email without the need to tab to the main warning.

An extended warning text preceding the email header as the one proposed above was particularly demanded by participants who had their screen readers on fast rate. Here, participant **P18** proposed adding a "*ping tone*, akin to the bell sound in [40], though not only *before* the warning, but also *after* it. In their view, the "*bookended*" warning alert will certainly get the attention of a blind user and make an "*immediate aural connection*" (**P7**) with the subsequent text of the warning. Here, participants like **P13** were aware that the alert should be "*a very short and somewhat unobtrusive sound, and also a configurable feature, so that [one] could enable/disable it or use your own sound*" as to prevent both non-adherence and habituation as a result of "*adding a second stream to the main aural stream of the screen reader*."

The interactive involvement of BLV users, past just presetting aural stream(s) as warnings, was also brought up as a potential improvement of the current elements. Making the banner warnings "*interactive elements*" with "*apprehensive text*" (**P3**) was seen as an improvement on the passive side of accessibility/usability so our participants recommended an active side too. Active for them was creating action elements

that were not just driven towards "*training the Gmail/Outlook filter*" (**P21**) (like the current "Looks safe," "Report not spam," and "It's not junk") but also helping them avoid both an immediate risk and the tediousness of future corrections. Here, participants felt it would be beneficial to have actions such as "*block this sender/subject*" (**P21**), "*delete email*," (**P3**), or "*move email to inbox*" (**P14**).

The low-vision participants, concordant with the recommendations and the ones noted in [30], stated that they would also benefit from a redesign that includes a "*relatable and comprehensible message*" (**P10**) in the banner text. Here, participants welcomed Gmail's choice to use distinguishing colors though they suggested using a pallet for colorblind users as the phishing warning banner, for example, comes in a stark red color. Participants also proposed an option to "*experimenting with a different shape than the rectangle*" (**P2**) so the captures the attention beyond the standardized angular design of email clients and the associated banners. This idea, in the view of participant **P14** who mentioned a new feature/shortcut of the JAWS screen reader called "smart glance," would equally help them quickly to "*areas that may stand out visually so it conveys those areas back to BLV users*."

5 Discussion

5.1 Accessibility and Usability Barriers

Our study aimed to uncover the accessibility and usability barriers in the engagement with the email banner warnings that BLV users experience under realistic conditions. These realistic conditions allowed participants to evaluate phishing, spam (Gmail), and junk (Outlook) emails with banners sent to their *own* emails instead of evaluating banners assigned to emails selected from phishing repositories [18, 40]. We also allowed participants to use their *own* device, their *own* client, and their *own* setup on assistive technology (including the speech rate or magnification ratio) instead of a setup where researchers control the device, the client, and the assistive technology for participation [18, 40]. As participants in reality might receive many emails with one type of banner warning (e.g., spam) and none of the others (e.g., phishing), we instigated a false positive email, that is a legitimate one that does not put participants at risk, to ensure they have both spam and phishing addressed to them. This allowed our participants to encounter email pretexts that pertain to their *own* email communication and not pretexts targeting mostly sighted users nor scenarios that might be unfamiliar to BLV users.

The Gmail and Outlook email banner warnings we evaluated in our study, designed for sighted users as visual frictions, did not account for the practice of many screen reader users to have a fast rate of text-to-speech translation, rendering the warning text in many cases of little use for them. In the Gmail phishing variant, the screen readers of our participants verbalized a "*phishing*" utterance that was assigned in addition to the warning. Appearing unexpectedly and sounding disjointed from the email (as it was followed by verbalization of the email header) it was also of little usability to our participants in deciding the nature of the email message ahead. The navigation was achieved with some effort, but participants would have preferred the banners (and utterances) to be implemented as intuitive interactive elements so they could better engage with the warning itself when sorting their email correspondence. The formatting of the banner warning text was also a barrier to our BLV participants because it confusingly asked them to nonetheless go through the email and decide if message the "looks safe," "report not spam," or indicate "it's not junk." This task, paired with the task of reverting incorrectly filtered emails, was deemed tedious by our participants and, in cases of 2-factor authentication emails, disadvantaged them to sighted users as it introduced a lot of timeouts.

Participants indicated that they do adhere to this narrow implementation of the email banner warnings, but nonetheless rely on their own judgment and use cues from the email header and body to decide how to proceed with a suspicious email. This judgment was based on aural or magnified "scanning" of the entire email [32], including the banners, which were scanned more so to *assess whether the warning banner was correctly applied to an email or not* (i.e., if an email was incorrectly labeled as spam). Here, our BLV participants enjoyed the advantage of avoiding phishing/spam emails that contained only attachments as their assistive technologies could not convey the content and thus led them to ignore these emails. As a strategy unique to BLV users, our participants pointed out cases of email pretext that included improbable or irrelevant requests contradicting their "human vulnerability" i.e., offering services that are offered mostly for sighted users.

Our participants didn't benefit much from the urgency of the warning because it left them wondering what was wrong with an email, what action to take, and whether the warning pertains to all of their email correspondence. For example, the Gmail phishing warning communicated danger, but the "basis" of that danger was unclear whether it related to "similar messages" the participant received or Gmail has noticed overall for all the Gmail users. The change of the warning sentence to a dialog with a question and answer for the Gmail spam banner variant was also confusing for our participants because of the absence of clear action of deleting the message instead of simply reporting it. Similarly for the Outlook wording, the inclusion of both junk/not-junk wording in subsequent sentences (with the latter actually being a link for "report not junk") and the lack of explicit reference as to who made the decision to "identify an email as junk" made it unclear how to act upon it.

5.2 Designing for Realistic Email Engagement

An adaptation towards usable accessibility when it comes to email banner warnings was successfully tested in [40] with

BLV participants. The implementation consisted of a bell sound, followed by a short introduction speech "warning! for details, press a shortcut key," which conveyed an overlay text of a general Gmail warning that urges users to be careful with the email message. This design, overall, achieved higher noticeability than the current Gmail designs. Participants who evaluated this adaptation also proposed redesigns that keep the audio warning, but instead of a shortcut key and an overlay text include just a warning text, either on the same page as the email header/body or replacing the email body. Some participants proposed using markers in the email list to help them distinguish ones needing further attention [40].

Our BLV participants based their recommendations relative to the interactive drawbacks they experienced through the tasks and reflections they performed in the study, as well as their past lived experiences, towards a "visceral" redesign of email banner warnings [31]. For example, participants proposed replacing the wording of the "*phishing*" utterance with a longer wording: "*Gmail marked this email as a phishing/spam; tab to see why*." This recommendation addresses for request for markers in the email list as an "early warning" proposed in [40] and might allow for higher noticeability of the warning (or we hypothesize so). The proposed adaptation of our participants allows for the full warning text to be implemented as either an overlay text, a text on the same page as the email header/body, or a text that replaces the email body. The linear tab action (and the equivalent action on a mobile device) could be configurable to allow users to choose the order where, in the order of email elements, the full warning will be read back to them (or not read at all).

Our participants also provided a design input to address the lack of comprehensibility and usability of the warning banners themselves, not just their *narrow* accessibility. Seeing "*why*" an email is marked as phishing, spam, or junk was deemed important for a BLV user as they can benefit from the banner warning not just for the email in question, but for other future emails, especially ones where Gmail/Outlook mistakenly apply (or omit) banners and put the messages in a wrong inbox. Transparently learning about automatic detection cues that they could also use, our participants felt that they could also benefit from nuanced formatting of the warning that helps them actively engage with emails. They proposed including a risk categorization level in the warning marker that would help BLV users "*quickly assess*" the email without the need to tab to the main warning.

The habits of using screen readers on "fast" were also considered in the design recommendations as participants saw the need for an alarm sound not just to notice the warning, but to grab and keep their attention. Here, our participants proposed adding a "*ping tone*" both before and after the warning marker. Aware of the disruptive nature of such an implementation (a concern also expressed by some of the participants in [40]), our participants thought it was also concerning from a habituation perspective as BLV users might increasingly ignore such a "*bookended*" warning alert after repeated exposure. To address this potential limitation, our participants proposed "*a very short and somewhat unobtrusive sound, and also a configurable feature, so that [one] could enable/disable it or use your own sound*."

Absent from the recommendations proposed in [40] was the redesign of the actions BLV users could take as part of the interaction with the warning. On this account, our participants requested creating action elements that were not just driven towards "*training the Gmail/Outlook filter*" for automated filtering (like the current "Looks safe," "Report not spam," and "It's not junk") but helping them avoid both an immediate risk and the tediousness of future corrections. For this, our participants proposed the inclusion of elements that allow to "*block this sender/subject*," "*delete email*," or "*move email to inbox*." Equally absent were the redesigns that cater to the needs of low-vision users that use screen magnifiers. Our low-vision participants contemplated a possible reformatting of the distinguishing banner colors to accommodate colorblind users with appropriate pallets as well as the use of "*different shapes than the rectangle*" for boxing the warning in order to achieve a higher noticeability beyond the standardized angular design of email clients and the associated banners.

5.3 Limitations

The realistic settings impose several limitations pertaining to our study. A limitation comes from the sample size that prevents generalization of the results to the entire population of BLV users, other email providers than Gmail and Outlook, and other devices than iOS iPhones, and Windows OS laptops. Another limitation is that we sampled English-speaking email users from the United States and used the email banner warnings in Gmail/Outlook in their English variant. Other language implementations of the warnings and the email client interface, as well as blind or low vision individuals from countries other than the US, might yield different results than ours. Aside from using only Gmail and Outlook (per our participant's own choice), a limitation comes from the fact that we used the implementation version of the respective banner warnings at the time of the study, that is, the first part of 2023. Future changes, adaptations, and improvements in the way Gmail and Outlook make their banner warnings accessible might render our results obsolete (which we sincerely hope will be the case soon). The current version of assistive technologies our participants used as part of the study could pose yet another limitation as any new features (e.g. advanced "smart glance") might transform how blind or low vision users access and use the email banner warnings, and with that, affect the overall findings.

A structural limitation is the choice of using the participants' spam/junk inbox instead of their own inbox. This research setup might have primed our participants with a hint that these emails were not to be trusted, even though

we took measures to minimize such a hint before participants did the tasks in our study. The priming, however, was a necessary compromise to introduce and test a novel methodology that brings experiments close to the real interaction with suspicious emails and away from laboratory settings (or settings that are not familiar or naturally occurring for participants, especially BLV users). We cannot fully generalize our results to BLV individuals' main inboxes because we were not permitted by our IRB to phish, spam, or tamper with Gmail/Outlook's filtering rules due to greater than the minimal risk to them. Another related limitation comes from the choice of the *false positive* email we used to initiate a phishing/junk warning. The particular sender and the email request itself might have contributed to a phishing tip-off in a greater capacity/effect than the warning itself. Other phishing emails might have yielded different results, but the only participant that received two phishing emails (**P17**) heeded the respective banner warnings in the same manner, giving at least a provisional validity to our findings in this respect.

Similarly, a limitation comes from the choice of spam/junk emails our participants arbitrarily selected in our study as any other message or an encounter with a legitimate email marked as spam/junk might have caused a different behavior around the banner warnings. Though we left our participants sufficient time and support to engage with the emails and the email banner warnings through the assistive technology of their choice, this might have not been insufficient for them to formulate a more informed expression about their overall suspiciousness assessment and ultimate decision about it. Despite all these limitations, and similar to qualitative studies in general, our study nonetheless provides rich accounts of BLV individual's' lived experiences with suspicious emails, which studies in laboratory settings hardly offer.

6 Conclusion

In this paper, we worked with 21 BLV email users to bridge the knowledge gap relative to how they experience, access, and use client warning banners for emails sent to their *own* email address. Our study reveals that BLV users access these banner warnings, but this access is not tailored to their needs, habits, and basic capabilities in realistic conditions. The non-intuitive usability of these banner warnings, thus, was experienced as a barrier to an equal security protection that sighted users enjoy by default. To address this exclusion, participants' own proposal was driven towards early warning alerts that include text and tones as well as comprehensible warning text that indicates the actual risk associated with each email that pertains specifically to them. We believe our results in realistic settings illuminate a compelling set of issues encountered by a population of email users dependent on assistive technologies, and as such, provide the locus for further inquiries that broadly consider the accessibility dimension of any inclusive human-centered security design.

References

[1] ABDOLRAHMANI, A., AND KUBER, R. Should i trust it when i cannot see it? credibility assessment for blind web users. In *Proceedings of the 18th International ACM SIGACCESS Conference on Computers and Accessibility* (New York, NY, USA, 2016), ASSETS '16, Association for Computing Machinery, p. 191–199.

[2] CLARKE, V., BRAUN, V., AND HAYFIELD, N. Thematic analysis. *Qualitative psychology: A practical guide to research methods 3* (2015), 222–248.

[3] COLES-KEMP, L. Inclusive security: Digital security meets web science. *Foundations and Trends® in Web Science 7*, 2 (2020), 88–241.

[4] COLES-KEMP, L., AND JENSEN, R. B. Accessing a new land: Designing for a social conceptualisation of access. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (New York, NY, USA, 2019), CHI '19, Association for Computing Machinery, p. 1–12.

[5] COLES-KEMP, L., ROBINSON, N., AND HEATH, C. P. R. Protecting the vulnerable: Dimensions of assisted digital access. *Proc. ACM Hum.-Comput. Interact. 6*, CSCW2 (nov 2022).

[6] CORNELL UNIVERSITY. Phish bowl. https://it.cornell.edu/phish-bowl.

[7] CYBERSECURITY AND INFRASTRUCURE SECURITY AGENCY (CISA). Avoiding social engineering and phishing attacks, 2021. https://www.cisa.gov/news-events/news/avoiding-social-engineering-and-phishing-attacks.

[8] DAS CHOWDHURY, P., DOMÍNGUEZ HERNÁNDEZ, A., RAMOKAPANE, K. M., AND RASHID, A. From utility to capability: A new paradigm to conceptualize and develop inclusive pets. In *Proceedings of the 2022 New Security Paradigms Workshop* (New York, NY, USA, 2023), NSPW '22, Association for Computing Machinery, p. 60–74.

[9] DISTLER, V., LENZINI, G., LALLEMAND, C., AND KOENIG, V. The framework of security-enhancing friction: How ux can help users behave more securely. In *Proceedings of the New Security Paradigms Workshop 2020* (New York, NY, USA, 2021), NSPW '20, Association for Computing Machinery, p. 45–58.

[10] FEREDAY, J., AND MUIR-COCHRANE, E. Demonstrating rigor using thematic analysis: A hybrid approach of inductive and deductive coding and theme development. *International journal of qualitative methods 5*, 1 (2006), 80–92.

[11] FRANZ, A., ZIMMERMANN, V., ALBRECHT, G., HARTWIG, K., REUTER, C., BENLIAN, A., AND VOGT, J. SoK: Still plenty of phish in the sea — a taxonomy of User-Oriented phishing interventions and avenues for future research. In *Seventeenth Symposium on Usable Privacy and Security (SOUPS 2021)* (Aug. 2021), USENIX Association, pp. 339–358.

[12] FURNELL, S., HELKALA, K., AND WOODS, N. Accessible authentication: Assessing the applicability for users with disabilities. *Computers & Security 113* (2022), 102561.

[13] GERBER, E. Surfing by ear: Usability concerns of computer users who are blind or visually impaired. *Access World 3*, 1 (2002), 38–43.

[14] GOOGLE. Advanced phishing and malware protection, 2023. https://support.google.com/a/answer/9157861?hl=en.

[15] HAYES, J., KAUSHIK, S., PRICE, C. E., AND WANG, Y. Cooperative privacy and security: Learning from people with visual impairments and their allies. In *Fifteenth Symposium on Usable Privacy and Security (SOUPS 2019)* (Santa Clara, CA, Aug. 2019), USENIX Association, pp. 1–20.

[16] HAYHOE, S., AND RAJAB, A. Ethical considerations of conducting ethnographic research in visually impaired communities. In *European Conference on Educational Research* (2000).

[17] KNIBBS, K. Twitter's layoffs are a blow to accessibility, 2022. https://www.wired.com/story/twitter-layoffs-accessibility/.

[18] LAU, E., AND PETERSON, Z. A research framework and initial study of browser security for the visually impaired. In *32nd USENIX Security Symposium (USENIX Security 23)* (Anaheim, CA, Aug. 2023), USENIX Association, pp. 4679–4696.

[19] LINDELL, J. Digital capital: A bourdieusian perspective on the digital divide. *European Journal of Communication 35*, 4 (2024/02/12 2020), 423–425.

[20] MICROSOFT. Aoverview of the junk email filter, 2023. https://support.microsoft.com/en-us/office/overview-of-the-junk-email-filter-5ae3ea8e-cf41-4fa0-b02a-3b96e21de089.

[21] MOSSANO, M., KULYK, O., BERENS, B. M., HÄUSSLER, E. M., AND VOLKAMER, M. Influence of url formatting on users' phishing url detection. In *Proceedings of the 2023 European Symposium on Usable Security* (New York, NY, USA, 2023), EuroUSEC '23, Association for Computing Machinery, p. 318–333.

[22] NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGIES (NIST). Glossary: Vulnerability, 2024. https://csrc.nist.gov/glossary/term/vulnerability.

[23] OPENPHISH. Openphish database. https://openphish.com/phishing_database.html.

[24] PETELKA, J., ZOU, Y., AND SCHAUB, F. 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* (New York, NY, USA, 2019), CHI '19, Association for Computing Machinery, p. 1–15.

[25] RENAUD, K. Accessible cyber security: the next frontier?

[26] RENAUD, K., AND COLES-KEMP, L. Accessible and inclusive cyber security: A nuanced and complex challenge. *SN Computer Science 3*, 5 (2022), 346.

[27] SASSE, A. Scaring and bullying people into security won't work. *IEEE Security & Privacy 13*, 3 (2015), 80–83.

[28] SELWYN, N. Reconsidering political and popular understandings of the digital divide. *New Media & Society 6*, 3 (2024/02/12 2004), 341–362.

[29] SHAREVSKI, F., AND ZEIDIEH, A. Designing and conducting usability research on social media misinformation with low vision or blind users. In *Proceedings of the 16th Cyber Security Experimentation and Test Workshop* (New York, NY, USA, 2023), CSET '23, Association for Computing Machinery, p. 75–81.

[30] SHAREVSKI, F., AND ZEIDIEH, A. "I Just Didn't Notice It:" Experiences with Misinformation Warnings on Social Media amongst Low Vision or Blind Users. In *New Security Paradigms Workshop* (New York, NY, USA, 2023), NSPW '23, Association for Computing Machinery, p. 10–23.

[31] STARK, L. The emotional context of information privacy. *The Information Society 32*, 1 (2016), 14–27.

[32] THEOFANOS, M. F., AND REDISH, J. G. Bridging the gap: Between accessibility and usability. *Interactions 10*, 6 (nov 2003), 36–51.

[33] VANCE, A., EARGLE, D., JENKINS, J. L., KIRWAN, C. B., AND ANDERSON, B. B. The fog of warnings: How non-essential notifications blur with security warnings. In *Fifteenth Symposium on Usable Privacy and Security (SOUPS 2019)* (Santa Clara, CA, Aug. 2019), USENIX Association, pp. 407–420.

[34] VERIZON. Data Breach Investigations Report. Tech. rep., Verizon, 2023.

[35] VIGO, M., AND HARPER, S. Coping tactics employed by visually disabled users on the web. *International Journal of Human-Computer Studies 71*, 11 (2013), 1013–1025.

[36] VOLKAMER, M., RENAUD, K., REINHEIMER, B., AND KUNZ, A. User experiences of torpedo: Tooltip-powered phishing email detection. *Computers & Security 71* (2017), 100–113.

[37] WANG, Y. The third wave? Inclusive privacy and security. In *Proceedings of the 2017 new security paradigms workshop* (2017), pp. 122–130.

[38] WASH, R. How experts detect phishing scam emails. *Proc. ACM Hum.-Comput. Interact. 4*, CSCW2 (oct 2020).

[39] WEB ACCESSIBILITY INITIATIVE (WAI). Accessibility, usability, and inclusion. https://www.w3.org/WAI/fundamentals/accessibility-usability-inclusion/.

[40] YU, Y., ASHOK, S., KAUSHIK, S., WANG, Y., AND WANG, G. Design and evaluation of inclusive email security indicators for people with visual impairments. In *2023 IEEE Symposium on Security and Privacy (SP)* (2023), pp. 2885–2902.

[41] ZURKO, M. E., AND SIMON, R. T. User-centered security. In *Proceedings of the 1996 Workshop on New Security Paradigms* (1996), pp. 27–33.

A Recruitment Email

From: Researcher's Email
Subject: Research Study Participation
Date:
To: Potential Participant

Hello,

I am contacting you as you are my personal acquaintance and you are a low vision or blind person under the legal definitions [see the end of the email]. I have recently participated in a research study that was looking to learn more about how people with visual disabilities utilize email client interfaces through their assistive technologies. The study was conducted as an audio-recorded Zoom interview during which I was asked to arbitrarily select an email from my spam folder, access the subject line and the sender, and provide my experience with the way the email was presented in my email client.

I was asked to share personal information, such as age, gender, ethnicity/race, education, email use, and visual disability diagnosis. The research study was anonymous and I had the opportunity to skip any question that I wasn't comfortable answering. The study took around 45 minutes to complete and I was compensated with a $25 Amazon eGift Card.

If you are also interested in participating in this research study, please contact [redacted] over email at [redacted] or telephone: [redacted].

Low vision or blind person that is legally blind is defined as: Anyone with acuity of 20/200 or field-of-view of 20 degrees or less in the better eye with correction; low vision with acuity up to 20/70 and field-of-view larger than 20 degrees in the better eye with correction.

Best, Participant

B Interview Script

Announcement

This interview is being audio-recorded for research purposes. You may stop the recording at any time. Do you consent to being audio-recorded? Recording starts now.

Questions and Tasks

1. How often do you see suspicious emails in your inbox?
2. How often do you see legitimate emails in your spam folder?

   Note: A baseline of "suspicious" and "legitimate" is established, based on the the participant's interpretation of what they think is suspicious (or legitimate). In case this interpretation differs from the definition of phishing/spam suspiciousness/legitimacy [7], a brief explanation of "suspicious" and/or "legitmate" is provided. This is done as such to minimize the probability of priming participants for the subsequent tasks.

3. For the purpose of the study, could you please check your spam or junk folder. Can you please go over a few emails of your choice — one by one — you have received in this folder. For each, just open it and let your screen reader access the subject and the sender (or read the subject and the sender with your magnifier).

   Carefully open each of these emails and just review the contents. We would like to ask some questions about this particular experience:

4. Have you noticed anything unusual about these emails? Please specify in as much detail as you can.
5. Have you noticed any warnings, banners, notifications, or labels about these emails? Please specify in as much detail as you can.

6. How do these warnings, banners, notifications, or labels assist you in determining the legitimacy of the email they were substantiated to?
7. Do you usually adhere and/or plan to adhere to these warnings, banners, notifications, or labels when determining the legitimacy of the email they were substantiated to?
8. (optional) Did you payed particular attention to each of the warnings, banners, notifications, or labels you encountered while accessing the emails for this study?
9. What cues do you usually use to assess the legitimacy of emails, in general?
10. Have you received any phishing, spam, or suspicious email training?
11. Have you ever been a victim of a successful phishing or spam campaign? If you are comfortable, please share your experiences from this event(s).
12. What would you recommend about how these email **banner warnings** that you have encountered in the study should be made adequately accessible for blind or low vision individuals or individuals who use assistive technology like screen readers and screen magnification?
13. Anything else you want to add on this topic or your experience with warnings about emails?
14. Demographic Questions [Gender, Race/Ethnicity, Age, Education, Visual Self-Identification, Device, Provider, Client, Assistive Technology, frequency of email use]

C Codebook

Email Banner Warning Accessibility and Usability

- **Accessibility** Codes pertaining to the *accessibility* of the email banner warnings shown in Figure 1, 2, or 3.
  - **Not Accessible** The participant expressed that email banner warning is not accessible
  - **Accessible** The participant expressed that email banner warning is accessible
  - **Accessible with an extra cognitive effort** The participant expressed that email banner warning is accessible, but requires an extra cognitive effort to make sense of the warning text
- **Usability** Codes pertaining to the *usability* of the email banner warnings shown in Figure 1, 2, or 3.
  - **Usable** The participant expressed that email banner warning is usable
  - **Unusable** The participant expressed that email banner warning is not usable
  - **Partially Unusable** The participant expressed that email banner warning is partially usable, but it should be improved further

Email Banner Warning Adherence and Attention

- **Adherence** Codes pertaining to the *adherence* to the email banner warnings shown in Figure 1, 2, or 3.
  - **Not Adherent** The participant expressed that they did not adhere to the email banner warning
  - **Adherent** The participant expressed that they did adhere to the email banner warning
  - **Selectively Adherent** The participant expressed that they did selectively adhere to the email banner warning
- **Attention** Codes pertaining to the *attention* to multiple email banner warnings shown in Figure 1, 2, or 3.
  - **Payed Full Attention** The participant expressed that they did payed full attention to the email banner warning (individually/multiple encounters)
  - **Payed Partial Attention** The participant expressed that they did payed partial attention to the email banner warning (individually/multiple encounters)
  - **Payed No Attention** The participant expressed that they did not payed attention to the email banner warning (individually/multiple encounters)

Email Assessment, Training, and Past Experience

- **Email Assessment** Codes pertaining to cues, criteria or rules of thumb used to determine a legitimacy of an email.
  - **Aural Cues** The participant expressed that relied on aural cues such as grammatical inconsistencies, typos, misspellings, out-of-order symbols
  - **Logical Cues** The participant expressed that relied on logical cues such as the improbability of an email request
  - **Elements in the Email** The participant expressed that relied on cues in the email structure such as the subject, email sender, timestamp, and body without attachments
- **Email Training** Codes pertaining to training about spotting and dealing with suspicious emails;

  - **Didn't Receive a Formal Training** The participant expressed that they did not received a formal training about spotting and dealing with suspicious emails, including phishing and spam
  - **Received a Formal Training** The participant expressed that they received a formal training about spotting and dealing with suspicious emails, including phishing and spam

- **Email Phishing/Spam Past Experience** Codes pertaining to instances of phishing/spam victimization in the past;
  - **Fell victim of a phish/spam** The participant expressed that they did fell a victim of a successful phish/spam in the past
  - **Didn't fell victim of a phish/spam** The participant expressed that they did not fell a victim of a successful phish/spam in the past

Accessibility and Usability Banner Warning Redesign

- **Accessibility Improvements** Codes pertaining to accessibility improvements
  - **Popup/Cover** The participant recommends the banner warnings to be implemented as pop-ups or overlays to give a sufficient non-visual friction for them before they access the suspicious email
  - **Warning Sound** The participant recommends the banner warnings to be augmented with a sound alarm in order to give a sufficient audio friction for them before they access the suspicious email

- **Usability Improvements** Codes pertaining to usability improvements
  - **Severity/Risk Level Indicators** The participant recommends for the banner warnings to include severity/risk level indicators to better discriminate between various levels of threats and risk exposures based on the email type

D Debriefing

Thank you for participating in our research on how users who are low vision or blind experience and utilize email warnings. This study aimed to examine whether people pay attention to warnings as a cue before they proceed to the website or not. The emails you selected might be phishing, scam, or spam so we suggest you proceed with caution and better delete them. So far, no research exists on how low vision or blind users are experiencing and utilizing emails warnings under realistic conditions. This is why we asked you to select and examine one email from your spam/junk folder. The AWS email you reviewed as instantiated by us, the researchers, and the email is harmless – that is, it is actually sent by AWS but Gmail/Outlook assigned the phishing banner warning nonetheless (we refer to this email as "*false positive*." You can safely ignore or delete this email. There is no negative consequence to you nor your past, present, or future involvement and use of the AWS services.

It was necessary for the researchers to withhold this information from you regarding the purpose of the study to ensure that your actions and answers to questions accurately reflected your behavior under realistic conditions. Your participation in the study is important in helping researchers identify the best ways to address the accessibility of the warnings by email providers that are equally usable and result in safe behavior specifically for blind or low vision users. You have the option to option to withdraw from the study after finding out about the mild deception now or no later than 30 days period after the data collection concluded. We will anonymize your the transcript of the interview, and to be able to remove your entry from the data bank of our research interviews you might need to provide descriptions of answers and reference to the spam/junk emails you accessed in your inbox so we can attempt to uniquely match your interview (something that only you know and no other person could use to identify you in the data bank).

The final results of this study will be published in a peer-reviewed journal or conference. Your results will not be available individually and your participation will remain confidential. If you have any additional inquiries please contact [redacted]. If you have questions about your rights as a research subject, you may contact [redacted] in the Office of Research Services at [redacted]. You may also contact [redacted] Office of Research Services if your questions, concerns, or complaints are not being answered by the research team, you cannot reach them, or you want to talk to someone besides them.