(Blind) Users Really Do Heed Aural Telephone Scam Warnings

Filipo Sharevski
School of Computing
DePaul University
Chicago, IL 60604
Email: fsharevs@depaul.edu

Jennifer Vander Loop
School of Computing
DePaul University
Chicago, IL 60604
Email: jvande27@depaul.edu

Bill Evans
School of Computing
DePaul University
Chicago, IL 60604
Email: wevans9@depaul.edu

Alexander Ponticello
CISPA Helmholtz Center for
Information Security
Germany
Email: alexander.ponticello@cispa.de

Abstract—**This paper reports on a study exploring how two groups of individuals, *legally blind* (*n*=36) and *sighted* ones (*n*=36), react to aural telephone scam warnings in naturalistic settings. As spoofing a CallerID is trivial, communicating the *context* of an incoming call instead offers a better possibility to warn a receiver about a potential scam. Usually, such warnings are visual in nature and fail to cater to users with visual disabilities. To address this exclusion, we developed an *aural* variant of telephone scam warnings and tested them in three conditions: baseline (no warning), short warning, and contextual warning that preceded the scam's content. We tested the two most common scam scenarios: *fraud* (interest rate reduction) and *identity theft* (social security number) by cold-calling participants and recording their actions, and debriefing and obtaining consent afterward. Only two participants "pressed one" as the scam demanded, both from the *legally blind* group that heard the contextual warning for the social security scenario. Upon close inspection, we learned that one of them did so because of accessibility issues with their screen reader and the other did so intentionally because the warning convinced them to waste the scammer's time, so they don't scam vulnerable people. Both the *legally blind* and *sighted* participants found the contextual warnings as powerful usable security cues that, together with STIR/SHAKEN indicators like `Scam Likely`, would provide robust protection against any type of scam. We also discussed the potential privacy implications of the contextual warnings and collected recommendations for usably accessible implementation.**

1. Introduction

Unwanted telephone calls inconvenience people on a daily basis because it's trivial for one to spoof Caller ID and to automate the process of calling a vast range of telephone numbers with relative ease [1]. Most of these calls attempt to "scam" people, i.e., defraud them financially or steal their identifying information, such as their Social Security number [2]. The Federal Trade Commission (FTC) shows the dire dimensions of this "inconvenience" as people in the US consistently incur a median loss of $1200, and millions of them reported their identity stolen through various telephone scam calls on a yearly basis since 2020 [3].

Enabled through communication technologies, telephone scams are not entirely unlike emails or SMSs that attempt to socially engineer a target. In all cases, the adversary employs a persuasive pretext with the goal of eliciting action from the target and attempts to maintain an impression of legitimacy. But telephone scams differ in that they entail targets deciding "on the spot" whether to take a call or not and, further, how to proceed. Protection-wise, thus, there is limited space for intervention. On an infrastructure level, an attempt has already been made to establish a call verification solution – STIR/SHAKEN – so a target call receiver could verify the authenticity of the Caller ID [4]. Usually, the target call receiver is "warned" either through a visual indicator displayed on the call screen or by the Caller ID being labeled as `Scam Likely`. The solution, however, is not perfect, and telephone call scammers find a way around it, or the calls are often times incorrectly labeled [5].

On a device level, users could use call-blocking apps (e.g., Hiya), but this protection is inefficient as it demands a proactive list update of unwanted numbers. Recently, new device-level protection was proposed by Google [6] where the *content* of an unwanted call is automatically screened using generative AI technology for known scam patterns. A visual warning – akin to those Google assigns for suspicious emails [7] (see Figure 1) – is then shown on the call screen to tell the receiver what the scam is about. Going beyond the call metadata and into the user call data, on a device level, is seen as necessary to effectively address the problem of spoofed Caller ID and append the STIR/SHAKEN protection against scams. Aware of the privacy implication such a solution might have (more details in 5.2) [8], Google envisions this to be an *opt-in* feature for future versions of their Android operating system [9].

Resembling Google's phishing or spam email banners, the telephone scam warning combines *visual elements* (e.g., red color shades, "End call" button, etc.) and *context* (e.g., "*Banks will never ask you to move money to keep it safe*") to cue targeted call receivers away from the scam "on the spot." While this intuitive approach might work for visually able people, past research shows that these warnings, as in the case of emails, are largely inaccessible for people with visual disabilities [10], [11]. People with statutory (legal) blindness are equal targets of telephone scams as their sighted counterparts [12], [13], so it is unknown whether contextual telephone scam warnings, such as Google's proposed solution, will have an accessible variant.

Google's telephone scam avoidance is not available yet for any user or, for that matter, accessibility testing [9]. It did nonetheless inspire us, meanwhile, to independently develop *aural* counterparts and test them in naturalistic settings (i.e., by cold-calling) with both legally blind and sighted telephone users in two variants (one *short* and one long, *contextual* warning) and two of the most prevalent real-world scamming scenarios (*interest rate reduction* and *Social Security* fraud). Based on the past evidence of visual call screen indicators' inaccessibility [12], [13], [14], [15], we reasoned that is better to proactively develop and test aural options that are by default accessible and ready for use – regardless of one's visual (dis)ability – when (and if) the new way of combating telephone scams becomes widely available. To these objectives, we conducted an empirical study to answer the following research questions:

- **RQ1:** How would *legally blind* people initially respond to an (a) *interest rate reduction* or (b) *Social Security fraud* scam call, compared to their *sighted* counterparts, in three conditions: (1) *without* a warning (baseline); (2) with a *short aural* warning; and (3) with a *contextual aural* warning?
- **RQ2:** How do (a) *legally blind* and (b) *sighted* people usually detect and deal with telephone scam calls?
- **RQ3:** What usability preferences do (a) *legally blind* and (b) *sighted* people have about telephone scam warnings?
- **RQ4:** What privacy concerns and design recommendations do (a) *legally blind* and (b) *sighted* people have about telephone scam aural warnings?

The key takeaways from our study are that the *aural* warnings in both short and contextual variants: (1) work, as only two people in the *legally blind* group "pressed one" and only did so either because of a problem with the screen reader or intentionally to waste the scammers time; (2) combined with the area code of the Caller ID, offer a robust cuing mechanism against telephone scam calls independently from the scam's pretext and fraud type, for example, financial or identity theft). The STIR/SHAKEN indicator (e.g., `Scam Likely`) does help fend off scam calls, but we found that it creates a problem where the screen readers don't verbalize the actual Caller ID in cases when the call is wrongfully labeled, leaving *legally blind* users without the opportunity to accept a legitimate call. Privacy equally mattered for our participants, with the lack of control and transparency being the topmost concerns. Design-wise, the aural warnings were well received, with few recommendations for haptic adjustments and text comprehensibility.

The main contributions of this paper are:

- Empirical evidence of the way *legally blind* and *sighted* individuals detect, screen, and respond to telephone scams in naturalistic settings involving aural telephone warnings;
- A novel, usably accessible telephone scam protection approach that cues receivers about the *context* or the scam's intent (fraud or identity theft);
- A set of privacy concerns as barriers for inferring the *context* of a scam call through AI means and associated design recommendations that consider both the accessibility needs and usability habits of individuals target of telephone scams.

2. Background

2.1. Telephone Scams

Receiving unsolicited calls with pre-recorded messages became an unavoidable burden for the vast majority of people in the US, despite efforts to curb their placement on both the infrastructural, regulatory, and user side [2]. Some of these calls are automated – colloquially referred to as "robocalls" – though not all of them are unsolicited or unwanted. For example, people regularly receive robocalls about their doctor appointment reminders, prescriptions, or political campaigns [4]. Some of these calls are spam, such as telemarketing, sales calls, or polls [16]. Some of these calls are *scams*, or unwanted ones aiming to exhort money or compromise the recipient's private information.

Telephone scams, usually coming from a spoofed caller ID, pose a unique problem because they could be automated and placed as "robocalls" and/or they could hide behind a seemingly innocuous spam pretext. Targeting the spoofing ID part, call authentication solutions such as STIR/SHAKEN have been introduced to help users "verify" numbers to assert a call is coming from a source it says it is [4]. To help with the automation and spamming parts, the FTC appended the curbing effort by introducing the national "Do Not Call Registry" where users can enter their number to be spared from unsolicited calls [16]. Phone companies, device manufacturers, and third-party application developers also offered call-blocking and call-labeling services so that users can receive as few unwanted calls as possible [17].

But even with such extensive and layered protection, telephone scams still reach users in alarming numbers. Just in the first half of 2024, the FTC received around half a million complaints about telephone scams that successfully scammed one in five users for a staggering $229 million in losses and more than 132,000 identities stolen [3]. Fundamentally, the problem with telephone scams is challenging to solve due to a couple of factors: (i) the nature of phone calls as a communication service; and (ii) the deception protection structure itself. While (scam) emails and SMSs are *non-real* time in nature – the recipient need not respond promptly – the telephone (scam) calls require the recipient to pick up the phone and respond in *real* time. One could "screen" the call and let it go to a voice mail, but the aforementioned FTC numbers suggest that that's not a common practice (besides, scammers tend to "pressure" recipients with repeated voice mail spam in order for them to pick up the phone [18]).

In other words, a recipient of a telephone scam must decide "on the spot" whether the call is unwanted or unsolicited in the first place and take action to terminate or proceed [2]. A recipient of a (scam) email or SMS, on the other hand, even if often urged by a pretext to respond "on the spot" (e.g., account termination, compromised credentials), could take advantage of the *non-real* time nature and look for "cues" of deception in the email/text or revisit the request at a later time before deciding what action to take. The transactional aspect of email and SMS communication thus allows for providers to apply extensive filtering in case of emails (e.g., "spam" or "junk" inboxes), reporting and user labeling (e.g., a "Report Junk" links at the bottom of any message from any unknown sender) [19], and elaborate warnings about the phishing, spam, or scam nature of the requests (e.g., banners reading "Why this message is dangerous" or "This message is reported junk") [7], [20].

The *real* time response limits, in the case of telephone scams, how much protection could introduce latency in the natural call flow in order to authenticate a caller and investigate the nature of the call for the purpose of blocking and/or labeling. There is no advantage of "spam" or "junk" call inboxes, any reporting could only be done after the call is answered, and the only warning that a user receives is a Caller ID labeled `Scam Likely` [21]. Telephone scam call recipients enjoy no advantage of images, logos, or any multimedia that usually helps email or SMS scam recipients cue a deception. Nor can they rely on skimming the text of the request to understand better what is being asked of them. If they let the phone go to voice mail and use a transcription service, the transcribed message is usually of such low quality that it is practically useless. Moreover, even if recipients might not trust a call (e.g., it comes from an unknown number or is labeled as `Scam Likely`), they might feel they have to answer it anyhow because it might be related to work, a personal engagement, or an expected callback.

2.2. Telephone Scam Prevention

Restricted as such, the recipients of telephone scams have mostly themselves to rely on for avoiding telephone scams as scammers rarely reuse phone numbers [22]. An independent empirical evaluation of the telephone scam problem run by Tu et al. revealed that scammers could effectively retrieve users' Social Security numbers by employing a government impersonation pretext and using a "neighbor spoofing" form of the caller ID (formatted such as it matches the first three or six digits of the recipient's phone in order to entice them to answer the phone call believing it is a neighbor or local organization, e.g., a school) [23]. In such a situation, Edwards et al. tested a STIR/SHAKEN verified indicator – a green circle containing the letter "V" and a text reading "Verified Number" as a label underneath the caller ID - to help users overcome the spoofed caller ID and local area code trick [4]. While this indicator helped users decide to trust the call, it did not help them determine the context of the call (i.e., the scamming intent of the caller), leaving the opportunity for scams to hide behind verified telemarketers or any other types of legal phone spams.

To determine the call context and spare the recipient from being bothered by the unwanted call in the first place, Pandit et. al proposed a virtual assistant that answers the call and interacts with the caller to determine the nature of the call [5]. A user evaluation suggested that users welcome additional information about the call context and an intermediary that emulates a "spam/junk" filtering functionality. As the implementation of such an intermediary might be costly and introduce prolonged latency in the normal call flow for users, Du et al. proposed an option where users could set up call-blocking policies using attribute-based policies (rather than the caller ID) that offer the flexibility of defining unwanted call based on predefined call contexts [1]. The solution relies on decoupled authentication and call setup processes, and it was only evaluated from a latency perspective; however, no user evaluation was presented about its usability and potential adoption.

Warning users about a potential scam telephone call while blending call-blocking functionality and implicit notification of the call's context is a feature promised by most of the third-party "anti-robocall" apps. To see whether this promise is kept, Sherman et al. evaluated the impact of interface design elements of the ten most popular "anti-robocall" apps on the user's "on the spot" decision-making [15]. Users involved in the evaluation positively selected the use of signage (e.g., general prohibition sign), background contrast, Caller ID verification, and the ability to monitor call logs or see app-provided statistics to determine the call context (e.g., "this phone number was blocked by X number of other users"). Building upon these findings, Munyaka et al. tested a variant of the call context determination element where the accuracy of various telephone scam call warnings was communicated to the users [13]. The results of this test indicated that knowing how accurate a telephone scam warning is affects the perceptions of the call context and nudges the recipient to err on the side of caution.

2.3. Telephone Scam Warnings Accessibility

A large-scale analysis of scam calls using a honeypot over almost a year revealed that telephone scams target vulnerable and at-risk populations, in particular older adults or recent immigrants [18]. A further evaluation showed that the at-risk populations are usually targeted with phone calls that use Social Security disability benefits as a pretext to engage with the call recipients [22]. Often, at-risk populations such as older adults, immigrants, and people with disabilities don't enjoy the same degree of protection when it comes to warnings about unwanted emails because the designs usually involve heavy graphical components, technical wording, or are not accessible at all [11], [24]. At the same time, these at-risk populations heavily rely on the telephone as a communication service partly because of the convenience and familiarity with it, partly because telephone calls are far more flexible and accessible than email or SMS (e.g., for people with visual disabilities).

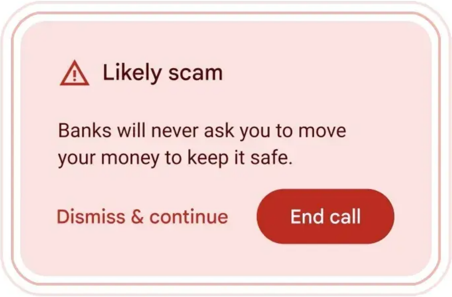
Recognizing that telephone scams place at-risk populations at a protection disadvantage, Sherman et al. also did an accessibility evaluation of the 56 "anti-robocall" apps with a set of legally blind users [12]. None of the apps were accessible through any assistive technology, offering no minimum color contrast for low vision users, no tags and labels on buttons for managing a call for navigation, nor any automatic audible alerts. Sherman et al. also involved legally blind users in the evaluation of aural warnings that spoke back variants of the verification label applied next or instead of the Caller ID [14]. Their results suggest that legally blind users favor plain language when communicating the type of call in order to be able to avoid confusion when deciding whether to take the call "on the spot" or not.

Munyaka et al. also involved legally blind users when testing the effect of telephone scam call warnings' accuracy on the ability to infer the context of the call [13]. Their findings indicate that visually impaired people heavily rely on Caller ID and the verification label (e.g., `Scam Likely`) to determine how to proceed with a call, and information about the accuracy adds a degree of confusion that could heighten their risk of being scammed (e.g., the blind users were confused why a Caller ID they recognized was also determined by the carrier to likely be a scam). The confusion was also a reason for Voight et al. to go beyond just the warning within the incoming call and propose a safe call-answering solution for older adults that involved an NFC card [25]. The idea is to build up a trusted source of callers to enable an accompanying card to give a warning to incoming calls from unfamiliar callers.

2.4. Contextual Telephone Scam Warnings

The protection measures against telephone scams, so far, hardly enable users to go beyond the Caller ID and determine the context relative to how to proceed. Restricted by the brief amount of time (e.g., several seconds, during the call ringing period), the design space for cuing the exact deception context "on the spot" is thus limited. To overcome this limitation, alternative scam protection is to alert users about the context *during* the call, i.e., after they answer it (or let it go to the voice mail, nonetheless, where the alert will also be recorded/transcribed). But such a solution requires *sampling* calls for conversation patterns commonly associated with scams and, using a generative AI engine, dynamically *insert* the warning before the call is connected.

Phone device manufacturers and app developers could feasibly implement this as an entirely client-site solution, but we haven't seen it yet for obvious risks of centralized control and privacy invasion [26]. Google recently hinted at a step towards such a solution, previewing a built-in Android OS feature that uses Gemini Nano (Google's generative AI engine) to dynamically offer a contextual scam cue "on the spot," even if the picked up [6]. An example contextual warning that Android plans to insert on the phone screens immediately after a call is answered is shown in Figure 1.



⚠ Likely scam

Banks will never ask you to move your money to keep it safe.

Figure 1: Contextual Telephone Scam Warning

Following Wogalter et al. [27] recommendations, the warning retains the cuing element signaling the nature of the call (Likely scam) though, in this case, it is not associated with the Caller ID like in the previous cases [4], [14]. This helps a receiver to identify the hazard (i.e., an unwanted call) "on the spot." Once this happens, the warning shows the context where the receiver is notified about the type and nature of the scam, for example, "*Banks will never ask you to move your money to keep safe.*" At this point, recipients are not directly offered an explanation of consequences if exposed to the hazard (i.e., loss of money), but are given just the context so they can infer the negative outcomes if they follow up with the call request (a general enough wording guarding against potential false positives and actual bank calls). The contextual cue comes as just-in-time and just-in-place warnings characteristic for (scam) emails and SMSs [28], [29]. Lastly, receivers are given directives for avoiding the hazard and offered to either "Dismiss & continue" with the call (an implicit button that requires clicking on the text) or click an explicit button that will "end call".

Obviously, the warning entirely relies on visual friction for a user to heed the context and avoid being scammed. A cursory accessibility evaluation similar to the one performed by Sharevski and Zeidieh for (scam) email warnings [11] would inevitably raise concerns about the presence of navigation elements so a screen reader could separately convey the banner, the context, and the button/action options. Following the concerns outlined by Sherman at al. [14] and Munyaka at al. [13], a visually impaired receiver could find it difficult to follow the context and intuitively make sense of its language as well as the wording of the buttons. And, specifically for legally blind people, there is the dimension of relevance to the accessibility. Even if Google makes the warning technically accessible, there is no guarantee that the warning will be relevant to a visually impaired person. For example, if the contextual warning is about an auto warranty expiration, the context might read "*Car insurers will never ask you to renew over the phone*," but driving a car is not something legally blind people usually do.

3. Study

As the regulatory push for curbing spam would necessitate some form of contextual telephone warnings to be conveyed to users [30], we wanted to explore how these "work" in an accessible variant, i.e., as *aural* warnings that pertain to both visually impaired and sighted users. We refrained from just an accessibility evaluation of Google's warning shown in Figure 1 because we were interested in how users – both legally blind and sighted – would react to relevant, aural contextual scam warnings in *naturalistic*, rather than in laboratory settings. We obtained an expedited Institutional Review Board (IRB) approval for a mild deception study that involved 36 legally blind[1] and 36 sighted adults from the US that have received scam calls in the past (the ethical considerations and research safeguards are elaborated in section 3.4). To obtain first-hand experiences with contextual scam warnings, we initially performed a pilot study to finalize the stimuli and adequately situate the main study into the naturalistic setting experienced by both legally blind and sighted people targeted by telephone scam calls.

3.1. Pilot Study

After we obtained consent, we created an aural translation of Google's contextual warning using both a feminine and masculine stock synthetic voice from the ElevenLabs library of synthetic voices [32]. As we could not exactly replicate the "Dismiss & continue" and "End Call" buttons, we experimented with an (1) "action" variant using the standard phone keypad buttons (press one to continue, press zero to end the call); and a (2) "narrative" variant where we simply asked the users to stay on the line or hang up using the standard "End Call" button. We conducted a pilot study with two legally blind and two sighted participants to test the aural and contextual variants (we obtained the initial IRB approval with the aural translation of the warning shown in Figure 1, and we amended the protocol later on with the final stimuli before we commenced with the main study). We presented the pilot participants with the most prevalent telephone scams per the FTC [33]. We then decided on the two most relevant real-world scenarios: one offered by Google [6] and associated with the FTC's example for *interest rate* scam and an additional one associated with *Social Security* to ensure the robustness of our findings.

During the pilot study, all the participants agreed to proceed with the second variant, where no action is asked of the receiver as that part of the warning might be mistakenly confused with part of the scam itself. Given that many legally blind people in the US receive Social Security benefits, our pilot participants suggest that we use the FTC's example for the *Social Security* scam, as all of them have received similar scam calls in the past. As the aural translation of the contextual warnings might run long, our pilot participants suggested we test a short warning variant that does not carry the context per se, but nonetheless urges the receivers to hang up on the call or proceed with caution. The resultant choice of contextual and the short warnings (preceding the scam message in the call flow) and the scam messages we used as stimuli, are given in Tables 1 and 2.

Through the pilot study deliberation, our legally blind participants suggested a slight rewording of the contextual warning to avoid the ambiguities present in the one proposed by Google [6], stating that "*This call is likely a scam*" (not just "*Likely scam*") and pointing to the scarcity principle of social engineering influence [34] by "*applying time pressure to make account changes*" instead of the just one resultant action – "*move money*." They suggested altering the tone between the warning and the scam voice. Accordingly, we used a feminine warning voice for the *interest rate reduction* scam (originally a masculine voice) and vice versa, a masculine voice for the *Social Security* scam.

TABLE 1: Interest Rate Scam Scenario Flow

Short Aural Warning	Contextual Aural Warning
This call is likely a scam. Hang up immediately, otherwise, proceed with caution.	This call is likely a scam. Banks will never time pressure you to make any change to your account over the phone. Hang up immediately, otherwise, proceed with caution.
Scam [33]	
Thanks to your good payment history and good credit score, you have been qualified finally for interest rate reduction between 0 to 5%. Several attempts were made to reach you. This is your final courtesy call before we are unable to lower your interest rates, so press one now.	

TABLE 2: Social Security Scam Scenario Flow

Short Aural Warning	Contextual Aural Warning
This call is likely a scam. Hang up immediately, otherwise, proceed with caution.	This call is likely a scam. Government agencies will never ask about your Social Security number or discuss legal matters over the phone. Hang up immediately, otherwise, proceed with caution.
Scam [33]	
We are calling you from the Department of Social Security Administration. The reason that you received this phone call from our department is to inform you that there is a legal enforcement action filed on your Social Security number for fraudulent activity, so when you get this message, kindly press one to connect with the next available officer.	

[1] Legally blind individuals in the US include those 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 [31].

3.2. Main Study Setup

Once we finalized the study stimuli shown in Tables 1 and 2, we structured a 2 x 3 x 2 study in naturalistic settings where two groups of participants (legally blind/sighted) were cold-called with one of the three possible study conditions (baseline, short warning, and contextual warning) in each of the real-world telephone scam cases (interest rate/Social Security). Situating the study in naturalistic settings, we had to ensure our participants were not exposed to a greater than minimal risk while we cold-called them so we could study the natural response to real-world telephone scam calls and obtain consent afterward. To do so, we first recruited two groups of participants, legally blind and sighted. For the legally blind group, we followed Gerber's advice when doing usability and accessibility studies with visually disabled people [35] and recruited individuals 18 years of age or older who have encountered unwanted calls in the past, use assistive technology, and are English-speaking and literate. We did a snowball sampling where we partially sampled personal acquaintances who are legally blind and partially a pool of legally blind participants that was recommended by one of our acquaintances. For the sighted group, we used the same inclusion criteria (sans the criterion for the use of assistive technologies), but we conducted a random sampling through Prolific.

Technically, we had six study conditions, and we reasonably balanced for a random assignment of six legally blind or sighted participants per condition (a total of 36 legally blind and 36 sighted participants). While our goal was to balance the conditions, we nonetheless made a particular effort to representatively sample each of the groups per the gender, race/ethnicity, and age dimensions (see section 3.3). We used a formal email approved by our IRB (see Appendix A) to approach each of the potential participants. The email invited our participants to join a 45-minute, audio-only, recorded Zoom interview session to discuss general experiences with unsolicited online communication. We used this deception pretext to remain general and not tip the participants off about the nature of the study. The only stipulation we had was that they use Calendly to sign up for the study and provide their phone number and email so we could send them reminders for their appointment.

We then used their phone number to cold-call them 24 hours before their participation time was scheduled (we initially built a buffer in the Calendly to enable us to do so on a rolling basis for all participants). We decided to do the call a day earlier to ensure participants do not suspect anything is amiss, or the call might be related to the study. Following the approach implemented by Tu et al. [23], we used CallFire to create six Interactive Voice Response (IVR) campaigns that were initiated from an innocuous phone number we purchased for the purpose of the study [36]. We settled to use one area code from a big metropolitan area in the US, as otherwise, we would have had to scale to at least 50 different innocuous phone numbers to match the sample's unique area codes, which was cost-prohibitive (we acknowledge this design compromise as a limitation). We set up the IVR logic to record whether a participant pressed the number one, which was indicated in the scams, allowing us to observe the natural action and the effect of the warnings (or their absence). Tu et al. [23] asked for the unwitting participants to enter four digits corresponding to the last four digits of their Social Security number, but we refrained from doing so because (i) our goal was to see if the participants would do as the original real-world scam requests (press one), and (ii) we deemed the collection of anything Social Security related as too risky for our legally blind participants who do receive Social Security benefits.

If a participant pressed one (or any other number, for that matter), they were automatically transferred back to a phone number controlled by the researchers where we had the ability to speak to them. At this point, the participants were debriefed (the entire debriefing about this mild deception is given in Appendix D), consent was obtained, and we asked them to join during their scheduled time a day later. If they did not do anything or let the call go to voice mail, we did the debriefing at the start of the scheduled interview and obtained consent for retaining their data from the study. We were aware that the classification of untrustworthy phone calls was predicated on the individual's telephone service provider, and we expected that we might encounter cases where the cold-call might not end up going through to the participant. We were also aware that we could not control for any STIR/SHAKEN Caller ID warnings (e.g., `Scam Likely`) presented by the participants' telephone service provider (or, for that matter, any app-based scam blocker, for example, Hiya) but we accepted these cases as part of the naturalistic settings.

3.3. Data Collection and Analysis

To collect data, we arranged audio-only recorded Zoom interviews on a rolling basis that lasted, on average, 45 minutes, and we compensated each participant with a $20 Amazon eGift Card ($1,520 in total, accounting for the compensation of the pilot participants too). Initially, the interview transcripts were not anonymized, but we removed any names and references to individual participants and deleted the audio recordings altogether (storing them on a secured server). Each interview was done with open-ended questions, listed in the interview script (see Appendix B). We concluded our recruitment with a sample of 72 (36 legally blind and 36 sighted) as we deemed a sufficient *information power* for answering our research question [37]. The demographics are given in Table 3.

Since we had to work with a degree of arbitrary selection of calls due to the natural settings, 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 [38], we first familiarized ourselves with the data, and then we completed a round of open coding for arbitrarily selected two interviews to capture the participants' actions and reflections. Then we discussed the individual coding schemes and

TABLE 3: Participant Demographic Distribution

	Legally Blind	Sighted
Visual Self Identification		
Totally Blind	18	0
Low Vision	18	0
Sighted	0	36
Gender Distribution		
Female	22	19
Male	13	17
Non-Binary	1	0
Race/Ethnicity		
White	22	24
Latinx	5	3
Asian	4	2
Black	2	4
Other	3	3
Age		
18-29	9	5
30-39	12	7
40-49	9	10
50-59	5	7
60+	1	7
Provider		
AT&T	3	6
Spectrum	2	2
T-Mobile	11	7
Verizon	10	15
Other	10	6

converged on an agreed codebook (see Appendix C). Two researchers used this codebook to individually code all of the interview transcripts. Whenever necessary, they added new codes for concepts not yet covered by the codebook (six in total). Afterward, we compared the results and resolved disagreements in the code assignments. We specifically did not include inter-rater-reliability calculations in our process, as our goal was to freely explore diverse topics associated with the natural behavior around telephone scam calls (codebook thematic analysis) [39]. We discussed the identified themes respective to our research questions and selected example quotations to represent our interpretations.

In reporting the results, we utilized verbatim quotation of the participants' answers as much as possible, emphasized in "*italics*" and with a reference to the participant as either **PV#**$_{\text{MW}}$ or [**PV#**$_{\text{MW}}$], where **P** denotes **participant**, **V** denotes the **vision ability level** of the participant (**B** - legally blind, **S** - sighted), and **#** denotes the order of participation. For each of the conditions, **M** denotes the **scam** message they received (**F** - Social Security fraud, **I** - interest rate reduction), and **W** denotes the **warning** they heard (**B** - baseline, **S** - short warning, **C** - contextual warning. For example, **PB2**$_{\text{FS}}$ denotes the second participant, who is legally blind and was cold-called using the Social Security fraud scam preceded by the short aural warning.

3.4. Trust and Ethical Considerations

As this was a study in naturalistic settings, it was important for us to balance the burdens and benefits of the cold-call as a mild deception relative to the Belmont Report [40]. To mitigate the brief violation of autonomy, during the debriefing, we acknowledged that the "burden" of receiving a cold-call was needed to test the *unbiased* utility of the contextual aural warnings towards dismissing the call as a scam or falling for it (press one on the phone's keypad, and subsequently entering a bank account or a Social Security Number). If we had invited the participants to test the warnings in laboratory settings, their attitudes or behaviors would have been *biased* by the fact that they had to hear the warning outside of a natural call flow, that is, the participants would have expected the call (a condition not present when a real scam call is received). Even if we called their *own* phone numbers in laboratory settings, they would have already known the nature of the call (fully disclosed in the process of obtaining the consent) and acted differently than they would in the normal course of their life where there are distractions and their focus is not entirely set on expecting a call (which would have been, if the study was done in laboratory settings and receiving a call is the primary task).

The *unbiased* utility of the warning is a "benefit" that would be enjoyed greatly among all the people with visual disabilities who depend on aural access to technology and warnings of any sort (phone calls, emails). Moreover, the legally blind population in the US would certainly benefit from contextual warnings, as they receive disability benefits, and scams often target recipients of such benefits (our Social Security Scam scenario) [41]. The benefits also pertain to the remaining population that does not have visual disabilities because phone calls are often received without actual access to a screen for a visual warning – for example, during driving, running, and calls received on earphones or external audio/microphone. For both the legally blind and sighted individuals, the "benefit" of the contextual warnings is the provision of the context itself as it explains why a call is a scam; an explanation that individuals could use it as an awareness measure to fend off future scam calls.

During the debriefing – either after accepting the cold-call or immediately after joining the scheduled interview appointment, participants were offered the option to *out-out*. None of them did. They empathically accepted our ethical consideration of "burdens" and "benefits," agreeing that the aural contextual warnings are of great personal benefit for their immediate family and friends. The *legally blind* participants, without exception, praised our work and felt a sense of contribution in a way that they were open to be cold-called to test and develop contextual warnings that would help all the blind people across the US, following the participatory principle of "nothing about us without us" involvement for people with disabilities [42]. As part of the debriefing, we offered the option to put participants' phone numbers to the national "Do Not Call" registry list [43] and helped them describe the process of reporting robocalls to

FTC. For our blind participants, we ensured they had the ability to fully access the registry website. We also offered resources to all participants on "How to Avoid a Scam" [44], which they found useful.

Only after we received the participants' permission that they were okay with proceeding with the study and discussing their experience with the cold-call, we commenced the interviews. We allowed them to verbalize the process, give comments, complaints, suggestions, and verbalize any other experience that was not necessarily with unwanted calls but other technologies (e.g., scam, spam, or phishing emails/SMSs) to allow for them to fully express the natural behavior "surrounding" their daily interactions with unsolicited communications. We offered emotional and accessibility support on an ongoing basis during the interview because talking about scam calls might be distressing. We thus allowed breaks if needed (recording stopped) or an option to reschedule the interview at a more convenient time.

The consent process made it clear that no names or addresses were collected, and their phone numbers were immediately removed from CallFire, Calendly, and our records after the participation was over. Participants were informed that they were free to stop and abandon any question at any point in time, remove any answers, or withdraw from the study up to 30 days after they participated. To attempt to locate and remove their data, we offered the participants to recall a unique way to identify their answers. No one requested removal after the participation. We also pointed out to our participants that they could act on the phone call from the study as they ultimately wished (e.g., delete a voice mail, add it to a scam call-blocking app, etc.). 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 participants that we were not involved with the way scam calls are handled or otherwise processed by telephone service providers, phone device manufacturers, or app developers. We also notified participants that we were not involved with the design and implementation of the contextual warning on which our study was based. 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 meaningfully *include* legally blind individuals in cases where contextual warnings are otherwise visually available for sighted people. We pointed out that, this goal, however, does not prevent from misusing our findings or misinterpreting them in making compromises for accessibility or removing such support altogether.

4. Results

4.1. RQ1a: Interest Rate Scenario

4.1.1. Initial Response. The initial response to the day-before cold-call for the *interest rate reduction* scenario (Table 1) is given in Table 4. We extracted call data records from CallFire to determine the action and call duration. None of the participants in either group "pressed one" as requested in the original text of the *interest rate reduction* scam. Overall, the participants in the *legally blind* group mostly let the call go to "voice mail, " hung up after listening to the warning, or simply hung up/declined. The *sighted* ones either let the call go to "voice mail" or hung up after listening to the warning in both variants.

TABLE 4: **Initial Response** – *Interest Rate* Scenario

Baseline	Short	Contextual
Legally Blind Participants		
Voice Mail	Hang Up*	Voice Mail
Hang Up	Voice Mail	Hang Up*
Voice Mail	Voice Mail	Voice Mail
Hang Up	Declined	Voice Mail
Declined	Hang Up*	Declined
Hang Up	Voice Mail	Hang Up*
Sighted Participants		
Voice Mail	Voice Mail	Voice Mail
Voice Mail	Hang Up*	Hang Up*
Voice Mail	Voice Mail	Hang Up*
Voice Mail	Hang Up*	Hang Up*
Voice Mail	Voice Mail	Hang Up*
Voice Mail	Voice Mail	Voice Mail
* indicates hanging up *after* listening to the warning		

4.1.2. Scam Event Recall. After the participants joined the interview, we debriefed them, obtained consent, and asked to recall how they acted on the cold-call they received 24 hours before. Table 5 breaks down the cues participants used to decide what action to take "on the spot." Per condition, the *legally blind* participants that answered but hung up decided to do so either because they did not know the number or they saw a STIR/SHAKEN indicator `Scam Likely`. This indicator was sufficient for one of them to directly decline the call. One of the *legally blind* participants that let the call go to voice mail said they do not check their inbox, and one of them said they listened to it, but the unknown number was a sufficient cue not to do anything. As all of the *sighted* participants let the call go to voice mail in the baseline condition, they decided not to act on it either because they did not know the number or heard the keyword "*press one*" which was a cue that this is a scam.

In the *short warning* scenario, the *legally blind* participants that picked up the call, heard the warning and decided to "*hang up immediately*" [**PB14$_{\text{IS}}$**] as the warning urged them to do so. The participants in this group that let it go to voice mail, decided not to do anything because "*the message said the call was likely a scam*" [**PB5$_{\text{IS}}$**]. One participant who declined did so because they did not recognize the number. The *sighted* participants that answered the call, equally, decided to hang up because the warning urged them to do so. Those who let the call go to voice mail did not take any action because either they didn't recognize the number or they saw an indicator `Scam Likely` on the recording.

In the *contextual warning scenario*, the *legally blind* participants that picked up the call, decided to hang up both

TABLE 5: **Action Recall** – *Interest Rate* Scenario

Baseline	Short	Contextual
Legally Blind Participants		
Unknown number	Warning, synth	Warning†, unkn
Unknown number	Warning†	Warning, synth
No Recall*	Warning†, unkn	Warning†
`Scam Likely`	Unknown number	`Scam Likely`
`Scam Likely`	Unknown number	Unknown number
Unknown number	Warning†	Warning
Sighted Participants		
Unknown Number	Unknown Number	`Scam Likely`
Unknown Number	Warning	Warning
Press One†	Unknown Number	Warning
Press One†	Warning	Warning
Unknown Number	`Scam Likely`	Warning
Unknown Number	Warning†, unkn	Warning†

* indicates *no* voice mail check or an immediate *deletion*
† indicates no action upon listening to the voice mail

TABLE 6: **Initial Response** – *Social Security* Scenario

Baseline	Short	Contextual
Legally Blind Participants		
Voice Mail	Declined	Voice Mail
Hang Up	Hang Up*	Voice Mail
Voice Mail	Voice Mail	Pressed One
Hang Up	Hang Up*	Hang Up*
Hang Up	Hang Up*	Voice Mail
Hang Up	Voice Mail	Pressed One
Sighted Participants		
Voice Mail	Hang Up*	Voice Mail
Voice Mail	Voice Mail	Voice Mail
Voice Mail	Hang Up*	Voice Mail
Declined	Voice Mail	Hang Up*
Voice Mail	Voice Mail	Voice Mail
Voice Mail	Voice Mail	Voice Mail

* indicates hanging up *after* listening to the warning

because the warning let them know the call was a scam and because they felt the actual scam message sounded "*synthetic*" or machine-generated (a robocall). Those who let it go to voice mail, upon inspection, decided to ignore the message because the warning said it was "*likely a scam*." The one participant that declined the call, did so because they did not recognize the number. All of the sighted participants who answered the call heeded the contextual warning and decided to hang up. One of those that let it go to voice mail said they did oot act on it because it had a `Scam Likely` indicator [**PS43$_{IC}$**] and one of them because the warning "*explained why they shouldn't*" [**PS45$_{IC}$**].

4.2. RQ1b: Social Security Scenario

4.2.1. Initial Response. The initial response, per the CallFire logs, to the day-before cold-call for the *Social Security* scenario (Table 2) is given in Table 6. Two participants in the *legally blind* group "pressed one" as requested in the scam, despite the presence of the long, contextual warning. Overall, the participants in this group mostly let the call go to "voice mail," hung up after listening to the warning, or simply hung up/declined. The *sighted* participants mostly let the call go to "voice mail", and in some cases, hung up after listening to the warning or declined the call.

4.2.2. Scam Event Recall. After the debriefing in the main interview, we set to learn more about the participants' initial actions, particularly the two *legally blind* participants who pressed one despite hearing the *contextual warning* (both of them were debriefed when they pressed one as the call was routed back to the researcher). Per Table 7, **PB18$_{FC}$** *accidentally* pressed one because of incorrect voice-over navigation, and the **PB9$_{FC}$** did so *deliberately*, reasoning:

> "*Someone was trying to stop me from pressing, you know, button number one (the warning). But I thought, I'm going to press it, I'm going to keep the person on the phone because if they're talking to me, they're not talking to some poor little old lady down the street who's going to give them their Social Security number*"

TABLE 7: **Action Recall** – *Social Security* Scenario

Baseline	Short	Contextual
Legally Blind Participants		
Synthetic Voice	`Scam Likely`	Synthetic Voice
Synthetic Voice	`Scam Likely`	Unknown Number
Unknown Number	Warning	Scam Sounding
Area Code	Warning	Area Code
Area Code	Area Code	Synthetic Voice
Unknown Number	Area Code	Warning
Sighted Participants		
Unknown Number	Warning	Warning†
`Scam Likely`	Unknown Number	Warning†
Unknown Number	Warning	Unknown Number
No Recall*	Unknown Number	Unknown Number
Scam Sounding	Area Code	`Scam Likely`
Area Code	Scam Sounding	Unknown Number

* indicates *no* voice mail check or an immediate *deletion*
† indicates no action upon listening to the voice mail

The remaining *legally blind* participants in the *contextual warning* scenario did not engage with the voice mail because the voice sounded synthetic or hung up after suspecting an unknown area code. The area code as a cue appeared in the *short warning* condition, too, as two of the *legally blind* participants were either suspicious of it and hung up or were reluctant to act on a voice mail. The remaining participants in this group hung up, heeding the warning, or avoided the call when they heard the STIR/SHAKEN label `Scam Likely`. The *sighted* participants that picked up the phone also heeded the warning, and those that let it go to voice mail suspected either an unknown number, a scam-sounding text, or a "scammy" area code.

In the baseline scenario, the *legally blind* participants hang up or let the call go to voice mail because they suspected an unknown area code, unknown number, or the message sounded synthetic. The *sighted* participant in the baseline scenario that declined the call had no recollection of it, and the remaining one let it go to voice mail because of a suspicious area code or unknown number, acting not because, upon inspection, the message sounded synthetic or was labeled as `Scam Likely`. The area code as a cue, interestingly, appeared in the *Social Security* scenario, perhaps because the scam usually targets people who collect Social Security/disability benefits. We noticed that the same pretext appeared in the work done by Munyaka et al. [13] and our participants, as did theirs, confirmed that "*a lot of Social Security junk comes from Florida*" so they always look suspicious for Florida area codes.

4.3. RQ2: Response and Telephone Scam Cues

When it comes to deciding whether people as targets of scams would take a call or not "on the spot," the reasoning is usually based on *cues of deception* that come in a couple of variants: (i) *call metadata and call presentation* and (ii) *plausibility context of the call* in the sense of reasons why one might get such a call. These two types of deception cues are exactly what emerged as patterns of meaning-making, out of which we developed two themes [37, p.113]:

(1) **Narrow Scam Cues** – focuses on deciding, "on the spot," whether an incoming call should be accepted and proceed with the request based on call metadata (known number, area code, STIR/SHAKEN indicators) and call presentation (voice familiarity, delays in speech);
(2) **Broader Scam Cues** – focuses on deciding, "on the spot," whether an incoming call should be accepted and proceed with the request based on the context, that is, the plausibility of receiving such a call.

4.3.1. Narrow Scam Cues. The first narrow scam cue the participants in both groups identified was whether the number calling them was known to them or not. Many stated that they never pick up calls from unfamiliar numbers, reasoning that "*all calls I need to receive are already in my contact list*" [**PB4$_{\mathbf{FB}}$**]. Some of the participants, like **PS49$_{\mathbf{IB}}$** noted that they went as far as to use a device feature on their iPhone to "*automatically silence calls from unknown numbers*" and consider picking up "*only if they see consecutive missed calls, as scammers usually are not that persistent and this might be a real call*". Some participants dismissed the unknown number as a narrow cue to scam because, as **PS71$_{\mathbf{IB}}$** pointed out, there was a broader context at play: "*a lot of times when you apply for jobs, jobs have agencies throughout different states and when an unknown number rings up, I actually want to pick so to make sure it's about a job that I applied for*" Others, like participant **PB25$_{\mathbf{IC}}$** explained that they would pick up calls from an unknown number, but would be looking for cues indicating an inauthentic conversation:

> "*Usually, you can tell if there's a long delay before someone starts talking when you pick up. Sometimes, if you just listen for about 3 or 4 seconds and there is no noise on the other side, you naturally say hello. Then you either start hearing a prerecorded thing or a bunch of people in the background from the call center. Then you know, well, this is certainly a bogus call.*"

For narrow scam cues during a call in progress, our participants looked at whether a recorded or artificial voice was used in the call. Not having an actual human on the other end of the line was heavily associated with phishing and scam calls, and participants stated that they recognized "*familiar voices used in such messages from TikTok*" [**PS4$_{\mathbf{IC}}$**]. Few *legally blind* participants declared they would immediately hang up when hearing inauthentic human speech. For example, **PB5$_{\mathbf{IS}}$** noted that "*sometimes you can hear they do not have a natural diction when speaking, sometimes you can hear a prerecorded message, and sometimes you can hear a call center with a background noise, but in all cases the voice gives out a fabricated quality*."

The area code was a second narrow cue that participants, both *legally blind* and *sighted*, referred to upon receiving an unsolicited call. For unknown phone numbers, participants would not pick up if they did not recognize the area code or the call was not a local one. Participant **PB10$_{\mathbf{FC}}$** reasoned a bit further where they would "*reject a call outside of their area call, but might call back if there are multiple missed calls*." That not always the area code might be a sole determinant of a scam call testified **PS70$_{\mathbf{IB}}$**, saying that "*people move out of state, sometimes even more than one state over time, and retain their area code so new local calls wouldn't match their own area code*" so deciding "on-the-spot" has to fall back on the broader context of a call.

Regarding the STIR/SHAKEN indicator, several *legally blind* and many *sighted* participants indicated that they never pick up calls that carry a "`Scam Likely`" designation. For sighted participants, deciding not to pick up on the spot because a call was labeled as a `Scam Likely`" was not just a single instance of fending off a scam call but more so about fending off future persistent targeting with unwanted calls. In the words of **P67$_{\mathbf{FB}}$**, "*The phone says the call is a scam and you might want to pick it up as it has a matching area code, but I know my friend picked it up once and they kept getting more and more calls, so I just ignore them.*" The evidence that *legally blind* users utilized this feature suggests that it is accessible using screen readers. However, some *legally blind* participants described cases where this method failed, either due to wrongfully labeled benign calls [**PB5$_{\mathbf{IS}}$**] or when a call is accepted before the label can be processed by the screen reader [**PB13$_{\mathbf{IB}}$**]:

> "*Honestly, I was not looking for the scam likely tag, because it said* `unknown` *number. And I was like, oh, that's my doc, so I didn't even get to the part where the voiceover would have told me. And it does read the whole phone number. Then it says, scam likely, so the idea of an audio (warning) just ahead of the incoming call is very intriguing.*"

4.3.2. Broader Scam Cues. When participants got deep enough into a call, they would assess information on the context they gathered. They demonstrated awareness of common scams and reported using that knowledge to detect fraudulent calls. For example, for participant **PB51$_{\mathbf{FB}}$** any request over the phone has to pass a really high bar of plausibility for one to proceed with the call because for "*social security, credit, payments, or anything money-related be sure, they're gonna get in touch with you one way or another and not just over the phone*." Several *legally blind sighted* participants mentioned that "*those agencies that need my personal information already have it, there is never a reason to disclose anything private over the phone*" [**PS70$_{\mathbf{IB}}$**]. A good deal of both *legally blind* and *sighted* participants explained nearly impossible situations such as *legally blind* people being contacted regarding their car insurance, as **PB16$_{\mathbf{IB}}$** pointed out:

> "*One that always gets me, because I'm blind, there would be no need for me to have car insurance because I kinda cannot drive, yet, really!*"

An approach of broader plausibility assessment emerged where participants in both groups would send calls to voice mail and inspect the transcript at a later time. Participant **PB62$_{\mathbf{FC}}$**, here noted that "*If it's important, they're gonna leave a voice message – the scam ones usually always convey a sense of threat of some sort*." The voice mail screening was also needed for elaborate scam schemes, like the one described by **PB28$_{\mathbf{IB}}$**:

> "*The voice mail said that in order to get my new Medicare card in time to have surgery which I needed, I had to give them my social security number and my Medicare information. Then, for this to actually happen, I also had to sign up to buy some back braces and knee braces that I didn't need, and that would expedite my Medicare card even further. So that's a cue*."

During the interviews, participants described one additional broader cue, namely calling back the alleged source of the call. They do so for any authority whose number is publicly known (e.g., the IRS) or familiar to them (e.g., their bank). For unknown sources, participants said they would *"google the number to see if it's anything legitimate"* [**PS65$_{\mathbf{IC}}$**] and decide on whether to call back if the call is "*specifically important to them*" [**PS66$_{\mathbf{IB}}$**]

4.4. RQ3: Usability of Contextual Warnings

To understand the unbiased utility of the contextual aural warnings, our participants had the option to recall the warning they heard, hear the alternatives, and inspect (aurally or visually) the one from Figure 1. There was an overarching consensus among the participants in all groups that the provision of *context* as a scam cue that combines the benefits of the narrow scam cues with a more informed "on the spot" decision that, ultimately, saves one's burden from searching for broader cues. From a usability perspective couple of themes emerged relative to the preference for context-based unwanted scam call protection:

(1) **Aural Warnings** – preference for *aurally* delivered warnings, "on the spot," and before the incoming call's content is received
(2) **Visual Warnings** – preference for an *visually* delivered warnings, "on the spot," and before the incoming call's content is received

The majority of the *legally blind* participants preferred both types of *aural* warnings, expectedly. There was a pronounced trend in preferences towards the contextual warning in the *Social Security* scenario as our participants felt it is particularly useful to give a context to a scam that "*comes close to the heart to blind people whose livelihood largely depends on Social Security benefits*" [**PB24$_{\mathbf{FC}}$**]. The majority of *sighted* participants, conversely, favored the visual warning from Figure 1. They did not dismiss the aural warnings altogether, though, and found the context useful, particularly the one warning them about the perils of discussing the Social Security number over the phone. The most prevalent justification for conveying context through a visual alert was the similarity with email or other similar alerts, that, in the words of the participants, "*it's right there in red and grabs your attention right away*" [**PS34$_{\mathbf{FB}}$**].

When probed to consider the warnings in a broader usability and accessibility context, all participants recalled scenarios where the aural warnings would have an advantage, for example, where one could configure "*an accent of their choice*," "*speed*," and "*volume*" of the delivery. They also saw the aural warnings as an excellent way to raise awareness about various scams as they emerge without the need to keep tabs themselves. In the words of **PS52$_{\mathbf{IS}}$**, the contextual aural warning:

> "*Goes along with the adage of like, 'give a man a fish he'll eat for a day; teach him to fish, he'll eat for the rest of his life'. So, the contextual warning is far better at educating people about what scams are there and what damage they can do to you*."

We found that the voice of the warning being used in the warning was perceived as a confounding factor in the protection against scams. Some stated that the masculine voice we used was too similar to those employed in scams, and as such, it might startle some people who are not yet used to it. Other participants, however, declared that it's good to grab the attention of the listener, is clear, loud, and *"gets the message across"* [**PB25$_{\mathbf{IC}}$**]. Also, some participants attested that the voices we used were appropriately authoritative. The female voice was described as calmer, which some preferred. It was perceived as *"trying to educate or warn you"* [**PB33$_{\mathbf{FC}}$**], whereas the male one was more related to solving a problem. One participant suggested recording their warnings in their own voice as a security mechanism, since scammers could not easily spoof that. Overall the voice was labeled *"super important"* and should convey *"the idea of trust and rapport"*. Especially for blind people, voice is crucial in authenticating the person or organization at the other end of the line, as further explained by **PB7$_{\mathbf{FB}}$**:

> "*As a blind person, you associate the voice with the brand, in perhaps the same way a sighted person would recognize most of the Chase bank branches look similar*."

4.5. RQ4: Privacy Concerns and Design Input

4.5.1. Adoption of Contextual Scam Warnings. Given the intrusiveness of the "screening" required to generate a contextual telephone scam warning, we asked all of our participants whether they would agree to such a service and what privacy concerns they might have. There was a broader segmentation of the sample, with some seeing privacy as an obstacle to the implementation of the scam warnings and others open to removing this obstacle under strict conditions. A couple of themes thus emerged relative to the adoption of contextual scam warnings:

(1) **Privacy above usable/accessible security** – provision of context is not acceptable as it will inevitably result in privacy violations;
(2) **Usable/accessible security with privacy safeguards** – provision of context is acceptable insofar privacy is strictly guaranteed.

Several of the *legally blind* participants in the *interest rate reduction* scenario were reluctant to use these warnings, for example, because "*it's neither legal nor preferred for one to listen to protected calls with clients*" [**PB17**$_{\mathbf{IB}}$]. The others were hesitant because they invoked negative experiences with "*targeted Facebook ads that clearly listen to [their] private conversations through the phone's microphone in real-time*" [**PB3**$_{\mathbf{IC}}$] and lack of transparency about "*who, how, and where, someone is going to use their data to train yet another AI engine*" [**PB25**$_{\mathbf{IC}}$].

The majority of the *legally blind* participants in the *interest rate reduction* scenario were fine using such a service, though they expressed concerns about potential privacy infringements. The main concern was the misuse of the terms of use and "*sending the data back to Google for 'improving the service'*" [**PB15**$_{\mathbf{IC}}$] as participants were open to giving feedback when the warning was wrongly assigned to a given call. **PB21**$_{\mathbf{IB}}$ mentioned that this "*will be just another part of [their] data to be collected*" as Google already provisions their Gmail, internet, and voice, and felt that this adds to the "*sense of entirely losing control over who tracks [their] data and for what purpose*." Those that had no privacy concerns were okay because "*[the warning] does help with just saving time screening calls*" [**PB6**$_{\mathbf{IC}}$].

In the *Social Security* scenario, three *legally blind* participants stated they would refrain from this client-site solution as they did not like the idea of someone else "*listening to their calls, even if it's for just a few milliseconds*" [**PB20**$_{\mathbf{FS}}$] or were afraid their "*data would end up on the dark web somewhere*" [**PB4**$_{\mathbf{FB}}$] Those that were okay to use such warnings were careful to point out that they would do so only on the condition that the data is not transferred outside of their device and have assurances that it would be "*deleted no later than 15 or 30 days*" [**PB33**$_{\mathbf{FC}}$] Those that welcomed the production of contextual warnings on the expense of sampling their conversations thought that this approach would, on a long run, help people "*stop getting these dumb calls for sure*" [**PB29**$_{\mathbf{FB}}$]. [**PB42**$_{\mathbf{FC}}$] appreciated the warnings as they "*felt more natural to them as a voice interrupts an unsolicited call and in a friendly manner basically asks you: Are you sure you want to do this?*"

All of the *sighted* participants in the *interest rate reduction* scenario were hesitant about the solution itself (not the warnings, per se) because they had a "*fraught sense of what AI is doing and how quickly it's taking a presence in a lot of spaces*" [**PS61**$_{\mathbf{IC}}$] Those that were open had similar reservations as "*it felt weird just knowing that there's an AI listening all the time*" [**PS57**$_{\mathbf{IS}}$]. The ones without privacy concerns reasoned that "*[their] carriers anyhow are selling [their] data to third parties that contribute to increase in scam calls so people have only their devices to protect themselves from incontinent scam calls*" [**PS72**$_{\mathbf{IC}}$].

Few participants in the *Social Security* scenario were reluctant due to negative experiences with other privacy-related issues, such as cookie consents. **PB68**$_{\mathbf{FS}}$ stated that they "*resent the abundance with which our system is driven by constant marketing and advertising, and really try to keep that noise to a minimum*" so they would decline any such an intrusion, even if it offered an obvious benefit. Those that were okay to opt-in but had reservations were mostly concerned about the false positives of the approach and the lack of transparency "*how AI would compensate for them on the expense of private information*" [**PB51**$_{\mathbf{FB}}$] The participant that welcomed the solution as-is, felt that it's much better option than the STIR/SHAKEN `Scam Likely`" because the contextual warning helps"*for the sake of safety and also not wasting time*" [**PB54**$_{\mathbf{FC}}$].

4.5.2. Design of Contextual Scam Warnings.

As the contextual warning, in both visual and aural variants, are still in an experimental phase (or at least, not mass rolled out as part of mobile OSs), we asked all of our participants to share ideas on how they would prefer being warned, cued, or alerted about incoming scam call "on the spot." Three design themes thus emerged relative to the way the contextual scam warnings might be implemented to best suit particular usability and accessibility needs:

(1) **Broadened accessibility towards usable security protection** – provision of *aural* warnings for any population regardless of their visual impairment;
(2) **Accessible security through customization and multidimensional altering** – ability for the *aural* warnings to be customizable relative to the type of voice, language, tone as well as inclusion of other non-visual alerts such as haptics – accent on *accessibility* as a means of reducing effort in fending off scam calls;
(3) **Usable security towards effort saving** – ability for the *visual* warnings to be customizable relative to known call interface affordances – accent on *familiarity and convenience* as a means of reducing effort in fending off scam calls;

An interesting theme emerged among both *legally blind* and *sighted* participants regarding the population that would benefit the most from the *aural* warnings in both the short and contextual variants. Participants felt that **older adults would particularly benefit from contextual aural warnings**. As one **PB14**$_{\mathbf{IS}}$ put it: "*they don't know all the scams, so the spoken context, without the need to do anything but just pick up the phone, would certainly help them*."

The *legally blind* part of our sample enjoyed the warnings and thematically saw a benefit for *customization* of the context of the warning as well as adding a *multi-dimensional alerting* for people with multiple disabilities. **PB11$_{\textbf{IS}}$** commented that it would be good for blind or low vision individuals "*to be able to select a language for the aural warning, for example, Spanish*" as there are many legally blind people in the US identifying as Hispanic. The blind/low vision participants also liked the alteration of the feminine/masculine tones between the warning and the scam text, but they felt an actual delimitation "*with one tone or two short consecutive tones*" [**PB3$_{\textbf{IC}}$**] would be more usable as that "*prepares them to approach the scam message itself more alerted*." Haptics were recommended frequently as an avenue of alerting people with not just visual disabilities but also deaf and hard of hearing or deafblind people. Having a "*default vibration pattern*" for regular calls and a "*preceding, different short vibration*" [**PB15$_{\textbf{IC}}$**] before the *aural* warning starts would beneficial for multiple participants, especially in busy, noisy environments like public transit.

The *sighted* part of our sample thematically saw a benefit for *effort saving*. Many of them mentioned this preference based on the "*convenience of not having to listen to the conversation*" [**PS39$_{\textbf{FB}}$**] and that they "*just wait for a short second to see if a warning screen will pop up*" [**PS46$_{\textbf{FS}}$**] to hang up straight away. But, the majority commented on the ambiguous context about the "*banks moving money*" as an action that is confusing and that they prefer to have the aural text variant instead (those that heard it). Few participants commented that the visual warning does not resemble the known functions of the green circular button for accept ("dismiss & continue") and the red one for hang up ("end call"), so, it is *confusing and not intuitive what action is the best to take here*" [**PS44$_{\textbf{FB}}$**]. The (un)familiarity gave a pause to a few participants who wondered whether people "*would effectively ignore the context and simply press the 'end call' button*" [**PS50$_{\textbf{IB}}$**] (e.g., by habituation). Here they saw the utility of having an *aural* warning from time to time so people *hear* the context and "*remain vigilant*" [**PS69$_{\textbf{FC}}$**].

5. Discussion

5.1. Usability and Accessibility Implications

Our results suggest a slight shift in the way people might respond to telephone scams in naturalistic settings, compared to the observations by Tu et al. [23]. Few years later, people still "hate" scam calls [2] and rely on the Caller ID to screen calls [23], though we saw increased screening by sending the call to "voice mail." This is reasonable to expect as the transcription got better, and the voice mail offers a buffer available for screening the call, particularly for legally blind users. We add to the evidence in [18], [22] that this at-risk population is, in fact, scam calls' targets, and they do welcome the idea of contextual warnings.

To this, our results confirm the findings in [4] that the STIR/SHAKEN indicator (e.g., `Scam Likely`) does help fend off at least some of the unwanted calls, even if spoofing of Caller ID has never subsided at all [3], [22]. The aural warnings we tested overcame the limitation of not having the context of the call identified by Edwards et al. [4]. Compared to the virtual assistant proposed in [5], the aural warnings eliminate the need for an intermediary and introduce a delay while retaining the agency of the receivers to ultimately decide how to proceed with the call. This is critically important for legally blind users as we found that the STIR/SHAKEN indicator (e.g., `Scam Likely`) creates a problem where the screen readers do not verbalize the actual Caller ID in cases when the call is wrongfully labeled, leaving them without the opportunity to accept a legitimate call. Of course, the contextual warnings are predicated on the screening of the actual call, but privacy concerns also remain in the case of the virtual assistant in [5].

Compared to the solution proposed by Du et al. [1], our results suggest that the contextual warnings, in the aural variant, could coexist with any call-blocking policies, especially for updating the list of predefined call contexts as the scams evolve over time. The same goes for any "anti-robocall" apps as they already implement some of the visual elements of the contextual warning shown in Figure 1, per the evaluation done by Sherman et al. [15]. We confirm the findings in [14] relative to the plain language of scam warnings demanded by blind or low vision users, and, our results show that added verbosity (context) imposes no accessibility barriers too. The wide acceptance of both the contextual and the short warning in our study suggests that they could also be complementary augmented with the idea of communicating the accuracy of the various telephone scam call warnings as suggested by Munyaka et al. [13] (though using a specific alter tone instead of signage).

5.2. Privacy Implications

The implementation of the aural warnings tested in our study is predicated on scanning the users' conversational data to infer the context in the first place [9]. Done through an AI engine, obviously, this approach raises alarms about the peril of privacy intrusions as well as censorship. The wide range of concerns expressed by our participants showed that usability and accessibility are important dimensions of scam protection, but not yet for the price of personal costs. The reluctance expressed by the participants is justifiable and relatable to negative experiences in the past with mass telephone surveillance in the US [45] and conjectures based on constantly growing evidence of data breaches [46]. A way forward, perhaps, might be to retain the AI approach but offer the option for telephone users to donate samples of their choice for the purpose of generating contextual warnings and an act of altruism [47]. We see this as a starting point, aided by privacy guarantees and full transparency of warning generation and provision processes.

5.3. Limitations

The naturalistic settings impose several limitations pertaining to our study. A limitation comes from the sample size, the telephone service providers used by our participants, and the top scams per the FTC [33] during 2024. We were limited to Google's concept of contextual warnings. Another limitation is that we tested the warnings in the native English variant to known US scams. One limitation is our choice to use only one area code for the cold-call. The cold-calling number matched the area code with 1 blind (*interest rate reduction*, **short warning**, Voice Mail) and 3 sighted participants (*interest rate reduction*, **short warning**, Voice Mail; *interest rate reduction*, **contextual warning**, Voice Mail; and *interest rate reduction*, **short warning**, Hang Up). For the rest, we used a non-matching area code, which prevents generalization of our results. The generalization is also limited by the low number of participants who actually answered the cold-call. We took the thematic analysis recommendation to tabulate the actions (Tables 4– 7) not as frequency counts towards claiming generalizability, but as a transparent depiction of the way participants responded and reasoned relative to the cold-call.

Another limitation is that we did not test the AI-based generation of the warning itself in various contexts. As our study was situated in naturalistic settings, we could not confine to only one device as participants used their *own* devices. Besides, the risk of privacy intrusion in our research settings would have been very high and might not justify the benefits of the study. A limitation comes from the synthetic voices we used to generate the warnings, and other "voices" might produce a different alerting effect than ours. We only went as far as the first step in the scam – the "press one" – but we cannot say with certainty whether the warnings we tested would help prevent people, legally blind or sighted, from giving up their details as the scam logic progresses. This was a necessary compromise to balance participants' privacy and well-being while introducing and testing a novel, more robust way of alerting people "on the spot" and on their *own* phone number. The naturalistic settings imposed a limitation to the degree to which the STIR/SHAKEN indicators were implemented, too, that is, we could not control for their presence or absence when we placed the cold-call.

6. Conclusion

Communicating the context of an incoming scam call does help fend off unwanted calls. Though an extensive experiment in naturalistic settings concerning two scam examples – interest rate and social security – we found that the *aural* way of warning the caller's intent was the preferred delivery for the legally blind participants, while their sighted counterparts perfected the *visual* just-in-time and just-in-place warning. These findings held even in the presence of Caller ID indicators or voice mail screening as alternatives to on-the-spot answering of a scam call. The implementation of these warnings is predicated on ensuring people's conversational data privacy, and we share a cautiously optimistic sentiment expressed by all the participants towards usably accessible protections against telephone scams of any sort.

References

[1] C. Du, H. Yu, Y. Xiao, Y. T. Hou, A. D. Keromytis, and W. Lou, "UCBlocker: Unwanted call blocking using anonymous authentication," in *32nd USENIX Security Symposium (USENIX Security 23)*. Anaheim, CA: USENIX Association, Aug. 2023, pp. 445–462. [Online]. Available: https://www.usenix.org/conference/usenixsecurity23/presentation/du

[2] H. Tu, A. Doupé, Z. Zhao, and G.-J. Ahn, "Sok: Everyone hates robocalls: A survey of techniques against telephone spam," in *2016 IEEE Symposium on Security and Privacy (SP)*, 2016, pp. 320–338.

[3] Federal Trade Commission (FTC), "Explore data," 2024. [Online]. Available: https://www.ftc.gov/news-events/data-visualizations/explore-data

[4] G. W. Edwards, M. J. Gonzales, and M. A. Sullivan, "Robocalling: Stirred and shaken! - an investigation of calling displays on trust and answer rates," in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, ser. CHI '20. New York, NY, USA: Association for Computing Machinery, 2020, p. 1–12. [Online]. Available: https://doi.org/10.1145/3313831.3376679

[5] S. Pandit, K. Sarker, R. Perdisci, M. Ahamad, and D. Yang, "Combating robocalls with phone virtual assistant mediated interaction," in *32nd USENIX Security Symposium (USENIX Security 23)*. Anaheim, CA: USENIX Association, Aug. 2023, pp. 463–479. [Online]. Available: https://www.usenix.org/conference/usenixsecurity23/presentation/pandit

[6] B. Heater, "Google will use gemini to detect scams during calls," 2024, https://techcrunch.com/2024/05/14/google-will-use-gemini-to-detect-scams-during-calls/.

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

[8] N. Lomas, "Google's call-scanning ai could dial up censorship by default, privacy experts warn," 2024, https://techcrunch.com/2024/05/15/googles-call-scanning-ai-could-dial-up-censorship-by-default-privacy-experts-warn/.

[9] Google, "Receive alerts for suspected scams during phone calls," 2024, https://blog.google/products/android/google-ai-android-update-io-2024/#scam-detection.

[10] Y. Yu, S. Ashok, S. Kaushik, Y. Wang, and G. Wang, "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.

[11] F. Sharevski and A. Zeidieh, "Assessing suspicious emails with banner warnings among blind and Low-Vision users in realistic settings," in *33rd USENIX Security Symposium (USENIX Security 24)*. Philadelphia, PA: USENIX Association, Aug. 2024, pp. 2083–2100. [Online]. Available: https://www.usenix.org/conference/usenixsecurity24/presentation/sharevski

[12] I. Sherman, D. A. Delgado, J. E. Gilbert, J. Ruiz, and P. Traynor, "Characterizing user comprehension in the stir/shaken anti-robocall standard," in *TPRC49: The 49th Research Conference on Communication, Information and Internet Policy*, 2021.

[13] I. N. Munyaka, D. A. Delgado, J. E. Gilbert, J. Ruiz, and P. Traynor, ""i used to live in florida": Exploring the impact of spam call warning accuracy on callee decision-making," in *2024 Usable Security and Privacy (USEC) Symposium*, 2024.

[14] I. N. Sherman, J. D. Bowers, L.-L. Laborde, J. E. Gilbert, J. Ruiz, and P. G. Traynor, "Truly visual caller id? an analysis of anti-robocall applications and their accessibility to visually impaired users," in *2020 IEEE International Symposium on Technology and Society (ISTAS)*, 2020, pp. 266–279.

[15] I. N. Sherman, J. D. Bowers, K. McNamara Jr, J. E. Gilbert, J. Ruiz, and P. Traynor, "Are you going to answer that? measuring user responses to anti-robocall application indicators." in *NDSS*, 2020.

[16] Federal Trade Commission (FTC), "National do-not-call registry faqs," 2024. [Online]. Available: https://consumer.ftc.gov/articles/national-do-not-call-registry-faqs

[17] ——, "How to block unwanted calls," 2024. [Online]. Available: https://consumer.ftc.gov/articles/how-block-unwanted-calls

[18] S. Prasad, E. Bouma-Sims, A. K. Mylappan, and B. Reaves, "Who's calling? characterizing robocalls through audio and metadata analysis," in *29th USENIX Security Symposium (USENIX Security 20)*. USENIX Association, Aug. 2020, pp. 397–414. [Online]. Available: https://www.usenix.org/conference/usenixsecurity20/presentation/prasad

[19] Apple, "Block, filter, and report messages on iphone," 2024, https://support.apple.com/guide/iphone/block-filter-and-report-messages-iph203ab0be4/ios.

[20] Microsoft, "Overview 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] J. McEachern and E. Burger, "How to shut down robocallers: The stir/shaken protocol will stop scammers from exploiting a caller id loophole," *IEEE Spectrum*, vol. 56, no. 12, pp. 46–52, 2019.

[22] S. Prasad, T. Dunlap, A. Ross, and B. Reaves, "Diving into robocall content with SnorCall," in *32nd USENIX Security Symposium (USENIX Security 23)*. Anaheim, CA: USENIX Association, Aug. 2023, pp. 427–444. [Online]. Available: https://www.usenix.org/conference/usenixsecurity23/presentation/prasad

[23] H. Tu, A. Doupé, Z. Zhao, and G.-J. Ahn, "Users really do answer telephone scams," in *28th USENIX Security Symposium (USENIX Security 19)*. Santa Clara, CA: USENIX Association, Aug. 2019, pp. 1327–1340. [Online]. Available: https://www.usenix.org/conference/usenixsecurity19/presentation/tu

[24] R. Bellini, E. Tseng, N. Warford, A. Daffalla, T. Matthews, S. Consolvo, J. P. Woelfer, P. Gage Kelley, M. L. Mazurek, D. Cuomo, N. Dell, and T. Ristenpart, "Sok: Safer digital-safety research involving at-risk users," in *2024 IEEE Symposium on Security and Privacy (SP)*, 2024, pp. 635–654.

[25] V. Voigt, R. Wiethe, C. Sassmann, M. Will, S. D. Rodriguez, and F. Alt, "Safe call: A tangible smartphone interface that supports safe and easy phone calls and contacts management for older people," in *Proceedings of the 22nd International Conference on Mobile and Ubiquitous Multimedia*, ser. MUM '23. New York, NY, USA: Association for Computing Machinery, 2023, p. 562–564. [Online]. Available: https://doi.org/10.1145/3626705.3631878

[26] A. Chan, C. Ezell, M. Kaufmann, K. Wei, L. Hammond, H. Bradley, E. Bluemke, N. Rajkumar, D. Krueger, N. Kolt, L. Heim, and M. Anderljung, "Visibility into ai agents," in *Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency*, ser. FAccT '24. New York, NY, USA: Association for Computing Machinery, 2024, p. 958–973. [Online]. Available: https://doi.org/10.1145/3630106.3658948

[27] M. S. Wogalter, V. C. Conzola, and T. L. Smith-Jackson, "Research-based guidelines for warning design and evaluation," *Applied Ergonomics*, vol. 33, no. 3, pp. 219–230, 2002. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0003687002000091

[28] M. Volkamer, K. Renaud, B. Reinheimer, and A. Kunz, "User experiences of torpedo: Tooltip-powered phishing email detection," *Computers & Security*, vol. 71, pp. 100–113, 2017.

[29] J. Petelka, Y. Zou, and F. Schaub, "Put your warning where your link is: Improving and evaluating email phishing warnings," in *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, ser. CHI '19. New York, NY, USA: Association for Computing Machinery, 2019, p. 1–15. [Online]. Available: https://doi.org/10.1145/3290605.3300748

[30] US Congress, "Fraud and scam reduction act," 2022. [Online]. Available: https://www.congress.gov/bill/117th-congress/house-bill/1215

[31] Social Security Administration, "404.1581. meaning of blindness as defined in the law," 2023. [Online]. Available: https://www.ssa.gov/OP_Home/cfr20/404/404-1581.htm

[32] Eleven Labs, "Generative voice ai," 2023, https://elevenlabs.io.

[33] Federal Trade Commission (FTC), "Phone scams," 2024. [Online]. Available: https://consumer.ftc.gov/articles/phone-scams

[34] R. B. Cialdini, "The science of persuasion," *Scientific American*, vol. 284, no. 2, pp. 76–81, 2001.

[35] E. Gerber, "Surfing by ear: Usability concerns of computer users who are blind or visually impaired," *Access World*, vol. 3, no. 1, pp. 38–43, 2002.

[36] CallFire, "Voice call brodcasts," 2024. [Online]. Available: https://www.callfire.com

[37] V. Braun and V. Clarke, *Thematic Analysis: A Practical Guide*. SAGE Publications, 2021.

[38] V. Clarke, V. Braun, and N. Hayfield, "Thematic analysis," *Qualitative psychology: A practical guide to research methods*, vol. 3, pp. 222–248, 2015.

[39] A.-M. Ortloff, M. Fassl, A. Ponticello, F. Martius, A. Mertens, K. Krombholz, and M. Smith, "Different researchers, different results? analyzing the influence of researcher experience and data type during qualitative analysis of an interview and survey study on security advice," in *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, ser. CHI '23. New York, NY, USA: Association for Computing Machinery, 2023. [Online]. Available: https://doi.org/10.1145/3544548.3580766

[40] The National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, "The belmont report: Ethical principles and guidelines for the protection of human subjects of research," 1979. [Online]. Available: https://www.hhs.gov/ohrp/regulations-and-policy/belmont-report/read-the-belmont-report/index.html

[41] American Council of the Blind, "Beware of Social Security Scams," 2024, https://www.acb.org/here-and-there-17.

[42] R. Jurgens, "Nothing about us without us," 2005.

[43] Federal Trade Commission (FTC), "Do not call registry," 2024. [Online]. Available: https://www.donotcall.gov

[44] ——, "How to avoid a scam," 2024. [Online]. Available: https://consumer.ftc.gov/articles/how-avoid-scam

[45] T. Edgar, *Beyond Snowden: Privacy, Mass Surveillance, and the Struggle to Reform the NSA*. Rowman & Littlefield Publishers, 2017.

[46] P. Mayer, Y. Zou, F. Schaub, and A. J. Aviv, ""now i'm a bit angry:" individuals' awareness, perception, and responses to data breaches that affected them," in *30th USENIX Security Symposium (USENIX Security 21)*. USENIX Association, Aug. 2021, pp. 393–410. [Online]. Available: https://www.usenix.org/conference/usenixsecurity21/presentation/mayer

[47] R. Wang, R. De Viti, A. Dubey, and E. M. Redmiles, "The role of privacy guarantees in voluntary donation of private data for altruistic goals," *arXiv preprint arXiv:2407.03451*, 2024.

Appendix A.
Recruitment Email

Header
Subject: Research Study Participation: General experiences with unsolicited online communication
Body
My name is [redacted], a professor at [redacted]. I am contacting you about a possibility to participate in a research interview discussing general experiences with unsolicited online communication. If you agree to participate, you will be asked to join us remotely via Zoom or phone for a 45-minute interview where we will ask a list of predefined questions. If you participate in our research study will be compensated with a $20 Amazon gift card which will be sent via the same email (the email will notify you that you have $20 gift card that you can redeem on Amazon) you have provided for contact, following the research interview. If you are interested in participating, please use the following Calendly link to find and schedule a timeslot that best suits your availability. Once you choose a timeslot and confirm in Calendly, you will receive an email including details on how to join the research interview which will be hosted via Zoom or a phone call. [redacted] If you have any questions, concerns, or are unable to use the above Calendly link to schedule your interview, please feel free to reach out directly to me either via email, text or phone call. You can also respond directly to this email. My email is [redacted] and my cell phone number is [redacted]. Thank you for your time in reading this email.

Appendix B.
Interview Script

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.

1) You might have received a phone call in the past that included an unsolicited request in relation to your Social Security number, tax return, utility bill, online accounts, or bank accounts. Tell us more about how you experience and handle such voice calls?

   1.1 Have you noticed anything unusual about these phone calls? Please specify in as much details you can.

   1.2 Have you noticed any warnings, notifications, or labels about these calls (e.g. "Scam Likely" as a caller ID)? Please specify in as much details you can.

   1.3 How do you usually review and decide what to do with these calls?

   1.5 What cues do you usually use to assess the legitimacy of a phone call?

   1.6 How often you receive potentially deceptive voice calls?

   1.7 Have you ever used any phone manufacturer (e.g., Apple, Google, Samsung, etc.) or cellphone carrier (e.g., T-mobile, ATT, Verizon, etc.) feature to block, restrict, or screen deceptive voice calls?

2) You might have received a phone call in the past that contained a message warning you that the call *might be potential scam*, containing an unsolicited request in relation to your Social Security number, tax return, utility bill, online accounts, or bank accounts. Tell us more about how you experienced and handled this voice call?

   2.1 What was your experience with this voice warning? Please specify in as much details you can.

   2.2 Have you noticed any other warnings, notifications, or labels about these calls (e.g. "Scam Likely" as a caller ID) in addition to the voice warning? Please specify in as much details you can.

   2.3 Would this voice warning factor in the way you review and decide what to do with these calls? If so, could you provide us more details how? If not, could you tell us why not?

   2.4 What you be interested to use such a feature to warn you about the nature of a phone call before you connect to the call?

   2.5 These warnings are custom created by us as researchers, but for them to be implemented in practice, the phone manufacturer (e.g., Apple, Google, Samsung, etc.) or cellphone carrier (e.g., T-mobile, ATT, Verizon, etc.) might need to listen for conversation patterns commonly associated with deceptive calls to be able to use an AI and apply the warning in a dynamic fashion. How would you feel about such a practice?

3) What is your perspective on how your cellphone carriers handle deceptive phone calls?

4) Have you ever been a victim of a successful deceptive phone call(s)? If you are comfortable with, please share your experiences with this event(s). [What lessons you have learn from here and how this episode affected your way of dealing with potentially deceptive phone calls afterwards]

5) Have you seen any other types of deceptive campaigns delivered through voice message but over other types of communication than phones (e.g. WhatsApp, social media, etc.)?

6) What would you recommend about how these warnings be designed adequately accessible for blind people or people with visual impairments?

7) Anything else you want to add on this topic or your experience with warnings about deceptive voice calls?

8) Demographics: How old are you? What race/ethnicity do you identify as? What is your gender, visual diagnosis, education level, cellphone provider, experience with deceptive voice calls?

Appendix C. Codebook

C.1. Initial Response

Codes pertaining to the initial response to the study phone calls, some of which included contextual scam warnings.

1) **Action** Codes pertaining to the *action* taken by the participant when they received the study phone call.
   - **Answered, Pressed 1** Answered the phone and pressed 1
   - **Answered, hang up** Answered the phone but hung up after listening to the content
   - **Let it go to voice mail** Let the phone call go to voice mail and then access the recording/transcript
   - **Declined** Declined the incoming voice call
2) **Recall** Codes pertaining to the *recall* by the participant about recent experience with an unsolicited call.
   - **Yes, confirming the study call** Received a call and the call was the one initiated by the researchers (participants were debriefed at this point and consent for retaining their data was obtained)
   - **Yes, but not the call from the study** Received a call but the call was not the one initiated by the researchers
   - **No** Never received any unsolicited call
3) **Cues** Codes pertaining to the *cues* noticed by the participant about their recent experience with an unsolicited call, if they had any.
   - **Unknown Number** Noticed that call was from an unknown number
   - **Area Code** Noticed that the area code differs from the area code of their phone
   - **Caller ID: Spam or Scam Likely** Noticed that Caller ID was labeled as Spam or Scam Likely
   - **Synthetic Voice** Noticed that voice of the caller sounded synthetic
   - **Delay** Noticed that there was a delay between answering the call and the speech on the other side
   - **Warning** [Except the baseline groups] Noticed that a warning was communicated to them about the call

C.2. Reflection

Codes pertaining to the response upon reflection about the study phone calls, some of which included contextual scam warnings.

1) **Action upon Reflection** Codes pertaining to the *action* that participant might have taken upon the reflection on the cues relative to the study (or other unsolicited) phone call(s).
   - **Answered** Answered the phone and pressed 1
   - **Answered, listened to the content** Answered the phone but hung up after listening to the content
   - **Screened the voice mail** Let the phone call go to voice mail and then access the recording/transcript to screen the call
   - **Declined** Declined any incoming voice call they don't recognize
2) **Experience with Unsolicited Calls** Codes pertaining to the *experience (e.g. the frequency)* with unsolicited calls in general.
   - **No experience** Haven't received any calls
   - **Never answered them** Received unsolicited calls but they never answer them
   - **Yes, seldom** Receive unsolicited calls
   - **Yes, frequently** Frequently receive unsolicited calls
   - **Yes, all the time** Receive unsolicited calls all the time (once or multiple a day, every day)
   - **Unknown Number** Suspicious when a call comes from an unknown number
   - **Area Code** Suspicious when the area code of the incoming call differs from the area code of their number
   - **Caller ID: Spam or Scam Likely** Suspicious about a call if they notice that Caller ID was labeled as Spam or Scam Likely
   - **Recorded Voice** Suspicious when they noticed the voice on the line is recorded
   - **Synthetic Voice** Suspicious when they noticed the voice on the line is synthetically generated
   - **Delay** Noticed that there was a delay between answering the call and the speech on the other side
   - **Context** Suspicious when the call is about personal information or about an inapplicable scenario (a car insurance expiration for blind people, for example)

C.3. Contextual Telephone Scam Warnings

Codes pertaining to the reception of the scam warnings used in the study (note, for those participants in the baseline condition, we played both recordings during the interview, after debriefing them about the goal of the study)

a) **Usefulness** Codes pertaining to the *usefulness* of the warning in fending off scam calls
   - **Short warning useful** Found the short warning useful

- **Long warning useful** Found the long warning useful
- **Both short and long warning useful** Found both the short and the long warning useful
- **Short warning useful; long warning not** Found the short warning useful, but the long one not
- **Long warning useful; short warning not** Found the long warning useful, but the short one not

3) **Population** Codes pertaining to the *target populations* the would benefit the most from the contextual telephone scam warnings
   - **Older People** Recommended using the contextual telephone scam warnings for alerting older people
   - **Blind/Low Vision People** Recommended using the contextual telephone scam warnings for alerting blind/low vision people
4) **Design** Codes pertaining to recommended *designs* contextual telephone scam warnings
   - **Tones** Recommended using various *tones* in addition to the warning to alert people with visual disabilities
   - **Haptics** Recommended using various *haptic patterns* in addition to the warning to alert people with visual disabilities
5) **Preference** Codes pertaining to the *preference* of begin alerted about incoming scam call
   - **No warning** Don't want to be warned about potentially incoming scam call
   - **Short Warning** Prefer the short contextual telephone scam warning
   - **Long Warning** Prefer the long contextual telephone scam warning
   - **Visual Warning** Prefer a visual contextual telephone scam warning
6) **Use and Privacy Concerns** Codes pertaining to the *use and privacy concerns* about the implementation of contextual telephone scam warnings
   - **Opt-in, no privacy concerns** Would opt-in for contextual telephone scam warnings and have no concerns about any potential intrusions of their privacy
   - **Opt-in, but privacy concerns** Would opt-in for contextual telephone scam warnings but expressed concerns about any potential intrusions of their privacy
   - **Opt-out, privacy concerns** Would opt-out from using contextual telephone scam warnings and due to, in part, privacy concerns about any potential intrusions of their privacy

Appendix D. Debriefing

Thank you for participating in our research on how individuals who are low vision or blind experience telephone scam calls compared to individuals who are sighted. This study aimed to examine how individuals would react to warnings about a potential scam call. The call you received yesterday was initiated by us, the researchers. If you heard any warnings, these were also created by us and the scam message (if you were in the baseline group) that was included too, based on the most popular scams tracked by the FTC. The entire call was innocuous – any digit that was requested to be pressed was automatically transferred back to a phone number controlled by us the researchers where you had the ability either to speak to them or leave a voice message. So far, no research exists on how low vision or blind individuals, compared to sighted individuals, experience and potentially might utilize telephone scam contextual warnings in naturalistic settings, that is, without knowing *a priori* that they might receive such a call. This is why we asked you to sign up to the study with the phone and we placed a call 24 hours before the time slot you have selected for participation.

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 response and utilization of the short and contextual warnings we created. Your participation in the study is important in helping researchers identify the best ways to address the design, accessibility, and preferential formatting of the warnings assigned by phone providers or phone manufacturers before connecting you to a potentially scam call. You have the option to provide *a posteriori* consent so we retain your action and response from the call and continue with the interview. After we end this meeting and compensation is provided, then we cannot remove your information since we did not collect any personal information (we removed your phone number from our data records).

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. We do not keep, record, or collect any personal information such as telephone numbers. 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.

Appendix E. Meta-Review

The following meta-review was prepared by the program committee for the 2025 IEEE Symposium on Security and Privacy (S&P) as part of the review process as detailed in the call for papers.

E.1. Summary

This paper presents a study of legally blind users and sighted users on their reactions to an aural variant of telephone scam warnings. The results suggested that the two legally blind users followed the instructions of the scammer but for reasonable reasons. The paper proposed a new approach for scam protection that includes the contexts of the call. The results also suggested that aural warnings can help users identify scam calls.

E.2. Scientific Contributions

- Independent Confirmation of Important Results with Limited Prior Research
- Provides a Valuable Step Forward in an Established Field
- Creates a New Tool to Enable Future Science
- Addresses a Long-Known Issue

E.3. Reasons for Acceptance

1) Novel solutions to scam calls, which is a long-standing issue in the security community
2) The naturalistic setting provides strong ecological validity
3) Sound methodology and the large number of participants indicate significant efforts