Designing Accessible Audio Nudges for Voice Interfaces

Hira Jamshed
School of Information
University of Michigan
Ann Arbor, Michigan, USA
hiraj@umich.edu

Novia Nurain
School of Information
University of Michigan
Ann Arbor, Michigan, USA
nnurain@umich.edu

Robin N. Brewer
School of Information
University of Michigan
Ann Arbor, Michigan, USA
rnbrew@umich.edu

Abstract

Older adults (65+) increasingly use voice assistants for information-seeking, but experience challenges and uncertainty in assessing information quality due to limited visual cues. HCI researchers have primarily used nudging, subtle approaches to guide users towards better decision-making, in visual interfaces to mitigate online misinformation and facilitate critical thinking. Thus, we extend nudging to voice-based systems to help older adults alleviate uncertainty in voice-based searches. We evaluate four audio nudge prototypes (i.e., non-speech and speech-based) with older adults (n = 34). Findings show that speech nudges more effectively prompt critical reflection than non-speech nudges because they are more disruptive. We discuss the significance of these findings for designing accessible audio nudges, highlighting the tension between disruption and accessibility best practices. Further, we propose that effective audio nudges should be explanatory and interactive to help older adults mitigate information uncertainty and raise open questions for the community about designing reflective nudges.

CCS Concepts

• **Human-centered computing → Empirical studies in accessibility**; **Empirical studies in HCI**.

Keywords

Older Adults, Nudging, Voice Technologies, Information Uncertainty, Accessibility

ACM Reference Format:
Hira Jamshed, Novia Nurain, and Robin N. Brewer. 2025. Designing Accessible Audio Nudges for Voice Interfaces. In *CHI Conference on Human Factors in Computing Systems (CHI '25), April 26–May 01, 2025, Yokohama, Japan.* ACM, New York, NY, USA, 16 pages. https://doi.org/10.1145/3706598.3713563

1 Introduction

Voice assistants (VAs), such as smart speakers and voice-enabled chatbots, are becoming more pervasive for supporting information-seeking practices among older adults (65+) [18, 81]. Researchers have highlighted their utility for older adults, especially those with vision or motor disabilities, as voice-based interaction does not require a screen, keyboard input, or mouse navigation [76, 82]. However, older adults have raised concerns about using voice assistants in sensitive contexts [12, 16, 102]. For example, when searching for subjective and complex health information, older adults describe VA responses as lacking depth, comprehensiveness, and cues that can help them assess information quality [16, 18, 32]. Missing information quality cues can contribute to information uncertainty, where users are unsure of the information's accuracy, completeness, usefulness, and credibility [103]. Further, the risks associated with information uncertainty can be exacerbated in sensitive or high-risk contexts (e.g., health information seeking) where overtrusting or misinterpreting information could lead to potential harm or death [10, 76, 111]. Thus, there is a need to investigate ways to effectively support older adults in mitigating uncertainty when searching for high-risk information using voice assistants.

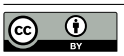
Prior HCI research has primarily relied on mis/disinformation detection algorithms and fact-checking tools to mitigate online information uncertainty [23, 54, 70, 78]. At the same time, researchers argue that such mechanisms often fail because users tend to rely on intuitive processing when consuming information [30, 63], especially when meaningful cues are absent, such as in voice-based systems [16]. However, there is limited to no research that examines how to facilitate reflective thinking and actively engage older adults in assessing information quality, including critically questioning the veracity of information sources [109]. Considering the urgent need to support older adults' decision-making process [16, 76, 111, 123], we take inspiration from prior work that has cited nudging as a potential approach to help users evaluate online information [1, 55, 62].

Nudges, defined as "*any aspect of the choice architecture that alters people's behavior in a predictable way without forbidding any option*" [107], have been designed by HCI researchers in many forms to encourage optimal decision-making [6, 22, 53]. While there are many nudge types, nudges promoting reflective thinking have helped individuals make informed decisions, and strengthened agency, autonomy, and independence [40, 105], values important for older adults [21, 74]. For example, presenting users with fact-checked articles next to potentially unreliable posts disrupts mindless scrolling and signals different perspectives [7, 22].

However, existing research has predominantly focused on designing nudges to mitigate online information uncertainty in visual interfaces, i.e., computers or smartphones [6]. Researchers and designers have yet to explore the design and evaluation of non-visual nudges, i.e., audio nudges, in voice-only interfaces for addressing older adults' information needs for voice-based technologies.

To address the gap, we investigate how older adults respond to *audio nudges* when searching for health information using VAs. Toward that goal, we designed four audio nudge prototypes (i.e., non-speech and speech-based) to address the following research questions:

(1) How do older adults respond to audio nudges?



CHI '25, Yokohama, Japan

ACM ISBN 979-8-4007-1394-1/25/04
https://doi.org/10.1145/3706598.3713563

(2) How do older adults envision audio nudges that can support critical voice information-seeking practices?

We conducted a remote Wizard of Oz [34] prototype evaluation study with 34 older adults (ages 65+), where participants listened to the VA's response for a specific health topic with and without audio nudges and provided feedback. Our findings show that older adults perceived speech nudges as more effective, even when falsely suggesting credibility despite inaccuracies in VA responses. However, older adults reasoned that speech nudges were more disruptive, prompting them to reflect critically on the credibility of the responses. They also shared their expectations for audio nudges, highlighting the need for a nudge that strikes a balance between encouraging critical reflection and guiding them toward obtaining credible information sources. Building on these findings, we reflect on practices for designing accessible and disruptive audio nudges. Additionally, we discuss the importance of designing reflective audio nudges that are *explanatory* and *interactive* to help older adults navigate and mitigate voice-based information uncertainty.

The work contributes to accessibility and aging research in the HCI community in the following ways. **First**, we provide an empirical understanding of older adults' reactions to audio nudges for voice-based information-seeking in high-risk contexts and present preliminary designs of audio nudges. **Second**, we discuss how disruptiveness is an essential characteristic of effective audio nudges while contrasting with accessibility research on ubiquitous audio design. **Third**, we raise open questions for the HCI community to advance audio-based nudging and discuss the potential of *explanations* and *interactivity* for designing reflective audio nudges, ultimately contributing to the discourse on helping older adults navigate voice-based information uncertainty.

2 Background

In this section, we summarize the literature on voice-based information uncertainty and the use of nudges for better information decision-making.

2.1 Voice Assistants and Information Uncertainty

Older adults commonly find voice-based systems, i.e., voice assistants (VAs), convenient and valuable because they are easy to use, seamlessly integrate into ongoing tasks, help overcome dexterity challenges, and foster independence [4, 12, 52, 111]. These characteristics have led older adults to use (or want to use) VAs for searching information online [76, 81]. The affinity towards VAs as information-seeking tools has called for increased scrutiny of trust in online information quality and such AI systems. For example, a systematic scoping review reports concerns about information quality and highlights how older adults are likely to encounter misinformation, conflicting information, and irrelevant information during their searches [122]. As a result, trust in AI systems varies among different older adults. For instance, in a three-week study with older adults using Amazon's Echo Dot, Pradhan et al. [81] found that some older adult participants preferred to verify the information from voice assistants using external sources, while others readily trusted the information they received. Prior research suggests that older adults' trust in AI systems is related to factors such as familiarity with the system, existing knowledge, and perceived humanness [41]. Shandilya et al. [95] emphasize the negative emotions (e.g., anxiety) and privacy concerns from unexplainable outputs of AI systems, often leading to confusion and low trust amongst older adults. Similarly, Weisz et al. [114] discuss the need to educate users to be skeptical of information quality, system inaccuracies, and biases/underrepresentation in training data. Relatedly, specific to VAs, prior research shows how older adults want the system to understand and respond to their detailed subjective queries but experience uncertainty, unreliability, and inconsistency [2, 12, 18, 94, 117].

In this paper, we focus on one of these challenges, *information uncertainty*, which occurs when users are unsure of the information's quality and use constructs like accuracy, completeness, usefulness, and credibility for decision-making [103] . The issue of information uncertainty becomes further aggravated in voice technologies because of the limited cues in voice-only interfaces. Voice interfaces have limited contextual elements that support users in judging information quality, like webpage design elements or consistent source information [65, 121]. Thus, prior work calls for better mechanisms to mitigate voice-based information uncertainty for older adults [16, 81]. In response to the call, we focus on health as a use case for studying information-searching contexts as it is common amongst older adults [111]. They often find VAs to be more cost-effective and less time-consuming than visiting the doctor, especially those with lower socioeconomic status [12, 76]. For instance, Zhang et al. [119] report that most older adults described "waiting several hours for just a few-minute consultation" as a reason for diverting to searching for information. The risks associated with health-related information uncertainty are high leading to potential harm or even death [10, 111]. This makes addressing information uncertainty in high-risk contexts like health an urgent problem.

Existing methods and strategies of assessing voice-based searches include judging response length and detail, repeating keywords, listening for the source, and having multiple rounds of conversation [16, 119]. Despite practicing these strategies, research also shows that users tend to ascribe human characteristics to VAs because of their human-like conversational abilities [80, 93]. For example, Brewer et al. [18] found that older adults expected the health information to mirror personalised advice from a medical professional, detailed responses instead of narration from websites, and follow-up questions. This expectation and uncertainty can often lead to detrimental consequences for older adults when qualitatively discerning reliable and applicable information. For instance, Brewer [16] investigated behaviors of voice-based information uncertainty and found that older adults engage in risky behaviors, such as trusting or using inaccurate information if it does not conflict with their existing knowledge. Considering how older adults find VAs to be a valuable, trustworthy, and convenient tool for information-seeking, this paper explores how we can better design voice systems to support older adults in mitigating information uncertainty.

2.2 Moving Towards Audio Nudges

Researchers have widely used nudging to facilitate positive behavior change [6, 22, 53, 107]. The HCI community has used nudges to facilitate, confront, deceive, influence, stop, encourage, or reinforce

behaviors or users' desires [6, 22]. While nudges affect people's choices, it is essential to note that nudges do not operate via manipulation. We use Hansen and Jesperson's [22, 40] conceptualization of a nudging technique that "prompts reflective choice," where users are engaged in a reflective mode of thinking and the intention of the nudge is transparent to the user, to study how nudges can effectively promote critical reflection. A reflective choice nudge engages users in deeper cognitive processing, motivates searching for additional information, and limits cursory decision-making [40].

Researchers have specifically called attention to the utility of nudging in supporting credibility judgments, yet these nudges are primarily designed for visual interfaces. For example, warning labels, fact-check alerts, color-coded content, webpage cues (e.g., engagement rate, number of references), alternative sources, or user-based accuracy classification are examples of visual nudges [8, 9, 45, 72, 77, 116, 120]. Specific to health information, researchers highlight the potential of nudging toward critical and reflective thought to combat misinformation [14, 18, 37, 61, 87, 123]. However, audio-based nudging remains largely unexplored [6, 53]. For example, research suggests that quickly assessing information quality via voice assistants is a critical challenge for older adults as there are no information quality cues for mitigating the uncertainty described in section 2.1 [12, 16]. As a result, older adults are left to create their own strategies with no means of confirming or verifying whether such strategies are effective [18]. Ineffective strategies can lead to undesirable consequences, such as practicing potentially harmful unconfirmed remedies or purchasing fraudulent health products [10, 16, 123].

Consequently, there is limited guidance on specific characteristics of effective audio nudges [6]. Our study utilizes and builds on prior work in the nudging and accessibility communities that explore different forms of speech in nudging and how sounds can convey complex visual information. Research related to verbal nudging has mainly involved understanding cognitive biases in conversational systems [3, 50], stimulating conversation [38], and encouraging healthy [20, 90] or pro-environmental behaviors [42, 73]. These studies demonstrate how nudges engage with users' biases, induce reflective thinking, enable further conversations, and push users towards healthy practices.

Relatedly, the accessibility community has researched innovative and creative ways of using non-speech sounds to support non-visual information access, mainly in the form of representative sound (auditory icons and musicons) or structured and abstract sound (earcons, spearcons, sonification) [35, 46, 49, 106]. Abstract non-speech sounds have specifically supported accessibility in data visualization [44, 96], image comprehension [108], programming [79, 92], and wayfinding and spatial perception [60]. Researchers argue that abstract non-speech sounds can efficiently present complex information, enable quick understanding of data, and be less obtrusive than synthetic speech [43, 69, 89]. For example, Siu et al. [100] find that adding sonification techniques to speech-based descriptions of data visualizations helped make non-visual data access more comprehensive, quicker to understand, and easier to verify the speech output. Similarly, Seo and Rogge [92] used earcons to mirror scanning through marked changes like additions, deletions, and modifications in code for blind programmers. Instead of listening to additional speech prefixes describing the change in every line, blind programmers could now hear the associated earcon in parallel with text-to-speech information from a screen reader and quickly get the context of changes. Such applications inspired our design of sound-based nudges to embody characteristics related to doubt and uncertainty. Additionally, research suggests non-speech alternatives for facilitating non-visual information access can be useful due to the low cognitive effort required, as humans can listen to and interpret complex signals involving different rhythms and harmonies and use sound to relate changes in dimensions to associated meanings [89, 101]. We refer to prior research that suggests loud, descending melodies, valley-shaped sounds, and minor mode patterns are significantly associated with low levels of trust [5, 88] to design the sound-based nudges.

We complement this literature by designing non-speech and speech-based nudges for signaling information uncertainty, gathering older adults' perceptions of doing so, and exploring nudges' role in improving voice-only systems.

3 Methods

We designed four audio nudge prototypes to understand older adults' responses to audio nudges (RQ1) and how they envision engaging with audio nudges in voice-only contexts (RQ2). We evaluated the nudges with older adult participants (n = 34). This section describes our nudge prototype design, study design, participant pool, and data analysis process.

3.1 Motivating Nudge Prototype Design

Prior work shows evidence of nudge's usefulness and effectiveness in visual contexts but not in audio contexts. As voice-based information-seeking becomes more popular, particularly amongst older and disabled communities, it will remain important to understand how to design audio nudges. For this exploratory study, we designed audio nudge prototypes to promote reflective thought and enable conscious choice based on prior work that shows that engaging the reflective system is important to enable conscious choice, motivate searching for additional information, and limit cursory decision-making [22, 40, 107]. However, as audio nudges have not yet been implemented in systems, it is unclear what form (e.g., speech or non-speech) is most effective in facilitating a user's ability to make informed decisions [6]. While speech can be more effective in directly signaling poor information quality, accessibility researchers show that non-speech sounds can also help effectively communicate information and reduce cognitive load [79, 97, 98, 100]. As such, we designed four audio nudge prototypes - two speech and two non-speech nudges (Table 1).

3.1.1 Speech Nudges. Informed by prior work on visual nudges that commonly consisted of a question or a statement [20, 90, 113], we designed two speech nudge prototypes to include a question ('Did you find this information believable?') and a statement ('The information I looked up online relates to a health topic'). The question nudge mirrors a *confrontational* nudge, intended to tap into a user's *regret aversion bias*, or the tendency to become a more careful decision-maker [86] and elicit critical thinking [6, 22, 53]. Similarly, by adding more information about the response, the statement nudge mirrors a *reinforcement* nudge. These nudges tend to

motivate certain behaviors (e.g., thinking about the source of the information) by increasing the behavior's salience in mind [6, 22, 53]. Researchers have shown that these types of nudges are effective in evoking a reflective response and facilitating critical thinking, which informed our speech nudge prototype design.

3.1.2 Non-Speech Nudges. Accessibility and sonification literature informed the sounds we chose for the two non-speech nudges. Because research describes how difficult it can be for users to connect abstract sounds to a specific purpose [11, 36], we chose sound patterns/wavelengths that previous work has shown to be distinct, such as how descending melodies are perceived to be less trustworthy [5, 88] . This can further elicit doubt or potential uncertainty [103]. For example, the *valley-shaped* sinewave in Fig. 1(a) consists of a decrease in frequency, forming a low point, followed by an increase in frequency. These frequency shifts correspond to variations in pitch, giving it a 'valley' effect and producing a characteristic dip in the auditory experience. The *descending* sinewave in 1(b) features a continuous reduction in frequency, creating a downward trajectory in pitch and giving the sound a 'falling' effect. We used sounds consisting of valley-shaped or descending sine waves as they are most likely to signal urgency or attention, signaling users to be more attentive to the information in the voice assistant's response [88]. Additionally, we mirrored the sound type on prior work evaluating accessible sounds [19, 83]. We used open-source audio-editing software [1] and created two sound nudges - a valley-shaped sine soundwave and a descending sine soundwave (Fig. 1).

3.1.3 Prototype Design. We recognize that there are an infinite amount of speech and non-speech nudge options, but we start with these initial four ideas informed by prior work to broadly probe for perceptions about audio nudges. We used the Voiceflow audio workflow platform to build and play each prototype across three health information-seeking scenarios - kidney stones, sinus infections, and migraines. We chose these scenarios as they are common health topics for older adults [66, 67, 84]. Similar to prior work [16, 17], we crafted responses to each scenario to mimic the structure, sound, and gender of commercial voice assistant responses and to elicit information uncertainty from participants. As such, the responses were intentionally vague or included incorrect information. Table 2 demonstrates different health scenarios and their corresponding VA responses. We pilot-tested these nudges and responses with a small group of older adults before using them in the interviews.

3.2 Study Design

We conducted a remote Wizard of Oz [34] prototype evaluation with participants using Zoom video conferencing, where participants reflected on the nudge's effectiveness and utility. We used pilot testing with a convenience sample of older and younger adults to iterate on nudge sounds and interview questions. First, participants were randomly assigned to a health topic (kidney stones, sinus infections, migraines) and engaged in a baseline task. In this baseline task, participants heard the VA's response to their assigned health topic without an audio nudge. In the next four tasks, participants heard the same response with audio nudge prototypes (presented in random order by the researcher). During each task, the researcher asked participants to rate each response's credibility after listening to the nudge (or no nudge in the baseline task) using a Likert scale of 1-5 (where 5 was the most credible). We probed for credibility as prior work has linked uncertainty, trust, and information credibility, where lower perceived credibility is associated with higher uncertainty [103]. We also asked participants about their expectations for interacting with nudges and perceived usefulness. After completing four tasks, we asked participants to rank the audio nudge prototypes from least to most favorite, explain their ranking, and describe how they envisioned future audio nudges (RQ2).

Finally, the researcher disclosed the purpose of the nudges after nudge rankings and asked participants to describe the nudge's effectiveness at signaling information quality. We intentionally disclosed this information at the end of the study to capture initial nudge perceptions (RQ1) without researcher bias. The researcher also read a statement at the end of the testing to caution participants against using any of the information presented in the responses.

3.3 Recruitment and Participants

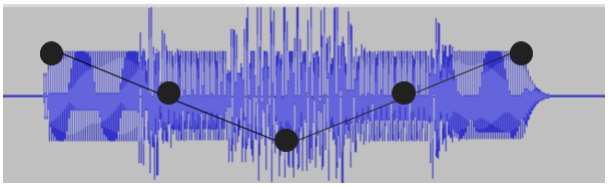
After IRB approval, we recruited participants via our university's health participant recruitment pool, a local participant research pool for older adults (anonymized), and several retirement and senior centers near a large Midwestern city in the United States. First, we asked potential participants to complete a screener survey about demographic information, aiming to recruit a diverse sample. From this survey, we recruited 34 participants (from 73 responses). Most participants were ages 65-74 years (n=23), with others aged 75-84 years (n=10), and 85+ (n=1). The rest of the details can be found in Table 3. Each participant received $25 by check or gift card after completing prototype testing.

3.4 Analysis

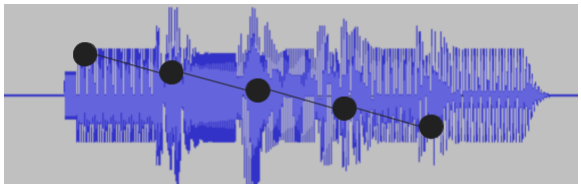
3.4.1 Qualitative analysis. After each prototype testing session, the lead interviewer engaged in memoing, reflecting on the session and participants' attitudes towards audio nudges. All interviews were audio-recorded. We auto-transcribed interview audio using Zoom, and the research team manually reviewed and corrected transcripts. We used a reflexive thematic analysis approach to analyze the collected data [15]. To do so, the first author read a sample of the transcripts (4-5 transcripts), conducted open coding, and generated an initial codebook using an inductive coding approach and codes that addressed proposed research questions. After discussing the initial codebook with the research team, the first author continued by iteratively coding each transcript and revising the codebook to capture emerging themes. Final themes included nudges as reflection, nudge perceptions (e.g., feedback, credibility, ending response, disruption), nudge behaviors (e.g., asking follow-up questions, searching elsewhere), envisioning future nudges, and response expectations (e.g., aggregating multiple sources). Although there were some themes related to evaluating voice assistant responses, we focused our findings on themes specific to nudges.

3.4.2 Quantitative Analysis. In the interviews, we asked participants to rate the voice assistant's credibility and to rank their preferred nudge, resulting in two forms of quantitative data. We use this

[1] https://www.audacityteam.org



(a) Valley-shaped sine wave. The line connected by black dots shows alternating changes in frequency, with a dip followed by a rise. This creates a distinct 'valley' effect, producing a distinct variation in pitch.



(b) Descending sine wave. The line connected by black dots shows a steady decrease in frequency, creating a downward trajectory in pitch.

Figure 1: An illustration of two different non-speech nudges' sine wave

Table 1: Examples of four different audio nudges

Nudge types	Audio files	Source URL
Statement nudge		https://dx.doi.org/10.7302/25208
Question nudge		https://dx.doi.org/10.7302/25208
Valley-shaped non-speech nudge		https://dx.doi.org/10.7302/25208
Descending non-speech nudge		https://dx.doi.org/10.7302/25208

Table 2: An overview of scenarios and VA responses (intentionally vague or incorrect [16]) used during the interviews

Health topic	Scenario	VA response
Sinus infection	You wake up feeling so not very good. You have a sore throat, cough, and pain in your sinuses. Thinking it might be a sinus infection, you want to take medicine, but you're not sure where to start. You ask your voice assistant for help.	According to Healthline: Use a nasal spray for at least 7 days to clear your sinuses. Taking over-the-counter antibiotics in combination will treat the infection.
Kidney stone	You recently started to eat more fruits and vegetables but begin to feel a slight pain in the left side of your lower abdomen. You want to know if your change in diet is impacting this discomfort and suspect that kidney stones are forming, so you ask your voice assistant about it.	According to Healthline: Eat plenty of foods that have low oxalate which can contribute to kidney stones. Eat fruits like bananas, blackberries, and blueberries and Vegetables like broccoli and cabbage.
Migraine	You are conducting your day as normal, but begin to feel a pain in your head. The pain increases to a pulsing sensation and you feel sensitive to bright lights and loud sounds. You know this isn't a typical headache and determine it must be a migraine.	According to Healthline: Caffeine in large amounts will help get rid of a migraine. Drink at least one cup every hour.

data to complement our qualitative data from the interview questions. We analyzed credibility ratings using non-parametric alternatives of repeated-measures ANOVA tests. We used non-parametric tests because (1) our data do not meet parametric test assumptions of normality and equal variance of errors, and (2) non-parametric tests are suitable for comparing more than two groups with small sample sizes and when the dependent variables are ordinal [29]. We used the Friedman test, which is appropriate for group responses in a within-subjects study design. We performed a pairwise comparison as the post hoc test using the Wilcoxon Signed Rank test with a Bonferroni correction. We used the Kruskal-Wallis test to assess whether the order of prototypes influenced credibility ratings, and the results showed no statistically detectable ordering effect. In addition, we fit a mixed-effects linear regression model, controlling for order while predicting the effect of the prototype on ratings. The results were consistent with the findings from the Wilcoxon Signed-Rank test (Table 4).

We analyzed nudge rank data using the rank-ordered probit (ROP) choice model to determine the expected ranking probability [110]. One participant did not complete ranking all four prototypes. We removed their data from our analysis. Among the participants,

Table 3: Demographics of the participants

		Participants (N=34)
Age	65 - 74	23 (67.6)
(N/%)	75 - 84	10 (29.4)
	85 - 94	1 (2.9)
Gender	Male	16 (47.1)
(N/%)	Female	18 (52.9)
Ethnicity	White	14 (41.2)
(N/%)	Black or African American	14 (41.2)
	Native American or Alaska Native	4 (11.8)
	Asian	2 (5.9)
Access to technology	Laptop/Desktop computer	30 (88.2)
(N/%)	Tablet computer	20 (58.8)
(yes responses)	Smartphone	33 (97.1)
	Voice assistant (Alexa, Siri, Google Assistant)	20 (58.8)

three participants provided two different rankings to a single prototype after remaining indifferent to it (e.g., ranking the prototype in the lower half of the range from least favorite to most favorite without distinguishing between 'second least favorite' or 'least favorite'). We chose to use the lower rank for those entries to model the ranking data. Taking the lower rank aligns with a more conservative modeling approach to mitigate overestimating user preference in the model. Moreover, we re-ran the model with order as an additional predictor to assess whether the model output was influenced by the prototype order. The results remained consistent (Table 5), indicating that order did not have a significant effect on the model's outcomes.

4 Findings

In this section, we describe the results from the quantitative analysis of participants' credibility ratings and ranking data in combination with the qualitative data from the prototype tests. In doing so, we detail how participants perceived (RQ1) and envisioned (RQ2) the audio nudges for conveying information uncertainty via voice-only prototypes.

4.1 Perceptions of Speech Nudges (RQ1)

We employed two types of speech-based nudge prototypes: a statement and a question nudge (Table 2). Participants described how both speech nudges helped with critically reflecting on the advice by reinforcing the VA's responses, but to varying degrees. We explored participants' *perceptions* of different nudges by asking them to rate how credible they found the responses after hearing each audio nudge prototype and a baseline response without an audio nudge. Table 4 illustrates participants' average credibility ratings for each audio nudge.

Based on our qualitative analysis, participants perceived the statement nudge to add more credibility to the VA response than the question nudge. However, our quantitative analysis did not uncover statistically significant differences in participants' credibility ratings between statement and question nudge (p-values for pairwise comparisons in Table 4).

After rating and listening to all the nudge prototypes, we asked participants to rank the four nudges from most (4) to least favorite (1) to capture their *preferences*. We employed the rank-ordered probit choice model to model the rankings. Based on the model, the expected probability of selecting the statement nudge prototype as the most favorite was higher compared to other alternatives and the difference is statistically significant (Table 5). Our analysis estimates that 63.9% of individuals would rank the statement nudge as their most preferred nudge.

4.1.1 Statement nudge falsely suggests credibility. Contrary to its purpose, most participants (n = 25) interpreted the purpose of the statement nudge ('The information I looked up online related to a health topic') to *add* credibility to the VA's response. For example, P29 said the statement, "*clearly points out the source of the information and tells you how credible the source is.*" They specifically focused on the nudge phrases "searched online" and "health topic" as giving more information about where the VA went to get the response and how it formed the response. P6 also pointed out how *"I don't remember if she [the VA] gave me the information about Healthline"* but that the statement suggested that *"it is a site and, apparently, the voice assistant went to the internet to look it up."* Similarly, P21 explained how the statement nudge *"clarifies, at least to a certain amount, what the source was even though it mentions Healthline."* P33 added to such sentiments by inferring that *"somebody has done the research, condensed it, and put it into a little phrase."*

Consequently, participants seemed more trusting of the VA response with the statement nudge and gave statement nudge high credibility ratings (mean = 3.35 and median = 3.5) compared to the other nudges (Table 4). For instance, P2 reasoned that *"In saying that you know this is where I got the information from, I will find that to be a little more credible because they're telling you where they got the information at least and that they did look it up."* Similarly, P0 considered the nudge as *"possibility that the advice may work"* because the VA *"did a little more research this time. She [the VA] is putting a little more into where she got the information from."* P18 increased their credibility rating from a 3 to a 4 (on a scale of 1 to 5, where 1 is the least credible and 5 is the most credible) because *"it backed it up. It gave credence to the fact that it wasn't just random."* P4 also thought the statement nudge, *"made it sound more credible by saying it looked it up."* Participants who were cautious with VA's in general also echoed similar sentiments, such as, *"I would be comfortable with the nasal spray but still not comfortable with the*

Table 4: Analysis of participant-provided credibility ratings (mean, median, and p-values) of different VA responses. Underlined p-values denote pairwise credibility rating differences are statistically significant at 95% confidence level.

VA responses	Without nudge	Valley-shaped nudge	Descending nudge	Statement nudge	Question nudge
Mean	3.56	3.16	3	3.35	3.15
Median	3.75	3	3	3.5	3
			p-values*		
Without nudge	-	0.18310	0.00856	1.00000	0.65810
Valley-shaped nudge	-	-	1.00000	1.00000	1.00000
Descending nudge	-	-	-	0.35340	1.00000
Statement nudge	-	-	-	-	1.00000

*with Bonferroni correction

Table 5: Analysis of participants' preferences for four different nudge prototypes. Underlined p-values* denote pairwise ranking differences are statistically significant at 95% confidence level.

Nudge prototypes	Valley-shaped nudge	Descending nudge	Statement nudge	Question nudge
Expected probability of ranking as the most favorite	0.0018572	0.1545506	**0.6393643**	0.2042386
		p-values*		
Valley-shaped nudge	-	0.062	<0.001	0.022
Descending nudge	-	-	0.001	1.000
Statement nudge	-	-	-	0.011

*with Bonferroni correction

antibiotics without contacting my physician. But I feel a little better because they did say they looked it up." (P2)

Even though the statement nudge added no content about the source or query topic, participants perceived the additional information in the statement nudge as *adding* credibility. The additional confidence in the VA's response also made the statement nudge a more preferred nudge. These quotes suggest that, even though the purpose of building reflective nudges in voice-only contexts is to motivate older adults to be critical of the information, nudges can also encourage them to overtrust the response by nudging in the wrong direction. We discuss the importance of including adequate yet subtle explanations in section 5.2 as it points to the complexity of determining information quality in voice-only contexts.

4.1.2 Question nudge amplifies doubt in VA. Several participants (n = 21) reacted to the question nudge ('Do you find this believable?') with strong skepticism. The nudge made them feel that the VA posed this question because it lacked confidence in its ability to answer their queries. This led participants to doubt the VA's ability to answer them and seek clarification elsewhere. For example, P3 felt that the VA was *"not a hundred percent sure of the advice it gave"* and answered the question nudge with *"Now that you say that, no."* When asked what they would do next, P3 claimed that "*the question would really make me Google!*" Similarly, P9 reasoned that the VA *"calls its validity into question"* because *"rather than saying 'did I answer your question?', it says more 'do you believe me?'"* This suspicion led many participants to say that they would *"quit the health assistant"* (P24) and *"get more information independently"* (P8). For example, P19 said, *"I would ask stuff that you know was fed like historical facts. Medical advice is much more serious. I prefer to look up like Mayo Clinic or WebMD or something like that, or check on my doctor for stuff like that."*

Consequently, participants critically reflected on their credibility ratings for question nudge. For example, P21 decreased the credibility rating from 2 to 1, explaining that *"in asking that question, the VA becomes even less credible. I mean, now, maybe this is not really great advice."* Others echo similar sentiments, e.g., *"I will be frank; it loses credibility because the person is going to look up more information. If she [the VA] is going to seek additional data, I am questioning why she did not have all the data when she began to issue the recommendations, you know?"* (P28) and *"I believe the responses through Siri and Alexa are pretty much vetted to be correct, but this makes me doubt a bit [...] The question could be a caveat that I should not accept it for the face or voice value, maybe I need to dig deeper into it"* (P22). Participants also reflected on the applicability of the advice after hearing the question. For instance, P29 pointed out, *"I mean, the response is believable, but it will make me doubt the credibility of the voice assistant because I believe it is meant to give me factual responses,"* and P12 changed their decision, *"I would think that I wouldn't run and get that coffee for my migraine like she suggested because if she's not sure, I wouldn't be sure either."*

These quotes suggest that the question nudge successfully signals participants to be cautious of the VA's responses by reflecting on whether the advice is *believable*. However, this also increases their

doubt in the VA's capability to answer their questions correctly and leads them to search for answers themselves by either contacting a medical professional or going to a trusted online source.

4.1.3 Question nudge as an interactive feedback mechanism. Several participants (n = 18) also interpreted the question nudge as a feedback mechanism for the VA. This led to speculations about how voice assistants form or verify the responses they give to the users. For example, P15 reflected how *"somewhere in the system, the devices program to check how accurately the recipient interprets or receives the information. So it's almost as if the system is trying to see if the information was accurate."* Similarly, P11 guessed that *"sometimes the VA listens to your responses to better inform it if people believe what it's saying"* and so will improve the response based on the feedback it receives.

Consequently, participants inferred the question nudge as an invitation to continue the conversation. For example, P14 said, *"Asking if it was believable leaves room for a question because it's not asking if it's helpful, it's asking, was it believable? And that leads to questions."* Participants also described how this interactivity affected their overall nudge ranking, with participants preferring the question nudge to the sound nudge prototypes. For instance, 21% of our participants selected the question nudge as the most favored, whereas only 3% of participants selected the non-speech nudge with the shape of a peak and valley (i.e., valley-shaped non-speech nudge). The difference in ranking is statistically significant (Table 5). As a potential reason, P11 shared that *"My favorite would be number 2 [the question nudge] because they gave the advice and then had a question for me to respond to,"* and P10 justified their preference for the question nudge by saying, *"Just the ability to get some feedback. Interactivity."* The participants further reacted to this feedback process positively, saying how *"it demonstrates a sense of communicating to the recipient that his or her feelings matter"* (P15). P19 further added how they *"like the fact that she didn't just give me information, just didn't say goodbye, or didn't say anything. The fact that she turned around and asked me a question made me feel better about the information that she gave me."* This opportunity for interaction also resulted in an increased credibility rating for these participants. For example, P2 explained the question nudge made the response *"seem a little more credible as they asked for my opinion on whether it was believable"* because now the VA will *"look into what it is actually responding with and make sure it gets a clear or more believable response."*

Overall, the question nudge prompted participants to reflect on the credibility and accuracy of VA's response and encouraged them to ask more questions from the VA. However, this happened indirectly, leading some participants to misinterpret the goal of the question nudge as a feedback mechanism for gathering accurate responses instead of signifying a lack of credibility in the response.

4.2 Perceptions of Non-Speech Nudges (RQ1)

Participants' reactions to the two non-speech nudge prototypes (i.e., valley-shaped non-speech nudge and descending non-speech nudge in table 1) were similar. As described below, participants perceived the non-speech nudges as less credible, leading to lower ratings for the VA responses in those prototypes. We also found a statistically significant difference between credibility ratings for baseline responses and responses of the descending non-speech nudge ($Z = 3.26$, $p < 0.05$, $r = 0.396$) (Table 4).

4.2.1 Ignoring Non-Speech Sounds. We found that most participants (n = 21) ignored the non-speech nudge played at the end of VA's response and disregarded the sound as an *"end of message sort of signal"* (P9). After probing, participants mentioned how they have learned to tune out sounds from technology because they are too ubiquitous in their physical and digital environments. For example, P28 laughed when adding to their interpretation, saying *"I mean, it makes no sense [laughs]. We're so accustomed to sounds nowadays that we begin to ignore them. So, I just ignore it, you know?"* Similarly, P5 pointed out how they *"didn't pay attention to it until you [the researcher] drew my attention to it, so I don't know if I would have paid attention."* P3 further compared it to their Alexa, saying how *"My Alexa often plays weird sounds at the end of a message, so I would think that is just another weird sound she plays. It is kind of an 'I wonder what that was for,' but then I don't really pay much attention to it and ignore it."* They also added how *"you get a lot of beeps and buzzes even on the phones, and sometimes you just don't know what they mean."*

Consequently, when asked to rate credibility and rank nudge prototypes from most to least favorite, these participants felt indifferent to the prototypes with non-speech nudges and ranked them lower than speech nudges (i.e., statement and question nudge). Based on our rank model, only 0.186% of individuals would select valley-shaped non-speech nudge as the highest rank (Table 5). After being told the purpose of non-speech nudges, participants pointed out that the non-speech nudges were ineffective in communicating that the information may not be credible. They echoed sentiments, such as *"it didn't say that to me at all. It was just extraneous noise, something for tech to edit on, and meaningless"* (P10) and *"they were just sounds; how am I supposed to take that as an interpretation of whatever she was trying to tell me? I just got nothing out of the sounds except that she was finished with the last one"* (P20).

For many of the participants, the sounds in the prototype were just another sound in their environment. Because of this, participants interpreted them as the VA communicating the end of the conversation and did not feel they helped in reflecting on the VA's response. This demonstrates how designing non-speech audio cues can be tricky because of how users have internalized the presence of sounds in their environment as not needing our attention. We return to this in section 5.1, where we discuss the conflicting needs between needing disruptive sounds as audio nudges and an accessible experience for communicating information uncertainty.

4.2.2 Disruption in the system. Another common interpretation of the non-speech nudges in the prototypes was that of a disruption in the VA system. When asked for their reasoning, these participants (n = 16) responded with concern or confusion and expressed dissatisfaction. Although we designed the non-speech nudges to be distinctive and represent that users should be cautious, P22 described the sounds as *"ominous"* and explained how *"it gave me a creepy feeling."* Similarly, P15 expressed annoyance and elaborated saying *"It was almost as if the system shut down prematurely. I really would question the information because of that sound. It was distorted, and I feel like there was more to say."* They then followed up by searching on Google to find more information and verify it. P17

related it to an alert and thought the VA took them someplace else, thinking that *"maybe it was a sound that directed me to some type of disclaimer on the information."* Participants even differentiated between the two sounds, like P2, who explained that descending non-speech nudge with a downscale sine wave *"felt like the VA got cut off as if there was something more to the information"* and that valley-shaped non-speech nudge with a peak and a valley sine wave *"felt more like an end."* The difference between the two types of non-speech nudges was reflected when participants ranked them. The descending non-speech nudge has a higher probability of being selected as the most favorite than the valley-shaped non-speech nudge (Table 5).

This group of participants felt the non-speech sounds communicated that the information might not be credible, suggesting an implicit understanding of the untrustworthy characteristics designed in the sound. For example, P22 added to their sound description above saying that the sound also left them with *"a feeling of a little bit of doubt about whether it was valid."* P26 reflected saying *"when they played that sound, I thought, oh, this sounds like second-rate advice, you know? So, in a sense, it did accomplish what it set out to do."*. Similarly, P30 gave the feedback that *"I think it did a very good job because I was not interested in the information immediately after I heard the sound. The sound makes it seem like the information is not finished. It sounds untrustworthy and it sounds unnecessary. It gives you that feeling."*

Overall, the participants expressed feelings of unease or doubt and conceptualized them as disruptions in the system. The quotes suggest that some participants are able to roughly detect the traces of low levels of trust in the sounds, but that this was infrequent.

4.3 Envisioning Nudges as Information Quality Cues (RQ2)

To understand how older adults imagine audio nudges as reflective and critical information quality cues, we gathered feedback on improving audio nudges. In reference to the speech nudges, participants spoke in-depth about the function and improvements to the design. However, in reference to non-speech nudges, participants expressed emotions elicited by the sounds and felt puzzled over the right improvements. As such, we differentiate the findings below accordingly and return to the implications in the discussion section.

4.3.1 Nudges Need to Guide. All participants stressed the importance of balancing the nudge's function of prompting critical reflection with guiding toward confirming information. Participants pointed to how the nudges caused them to overtrust the VA's response, question the VA's ability, or reflect on the information without encouraging or directing towards specific information about its credibility. To create the space for confirming and critiquing content, participants envisioned nudges to subtly highlight the credibility status of the current source and encourage the use of other sources by offering the option of asking for them. For example, P6 recommended that the nudge says something like, *"this might not be credible, and you can go to other sources to look up information,"* and asks the user, *"Would you like me to go to another source for you?"* Similarly, P5 pointed out that the nudges *"can be of a better service"* by *"serving as a conduit for further information"* and reminding users to *"not take any statement as a complete resolution or answer."*

As shown in table 1, the speech nudges were designed to be subtle and brief. While some participants understood the goal of the speech nudges, many gave examples of how the nudges could be further improved. For example, P28 stressed how the speech nudges would have been more useful if they had encouraged them *"to continue to pursue and gain secure information."* They suggested that the nudge could ask them to *"look further into what we [the user] could do"*, guiding them toward obtaining credible information instead of asking if the information was believable as it made them question the VA's ability instead of *"communicating dissemination of untrustworthy information."* P7 explained the statement nudge as their favorite compared to the other nudges because *"it did tell you at the end that it looked it up, but it finished after explaining where it came from. Just right to the point."* P23 focused on the length of the statement to elaborate on simplicity and thought that *"it hit just about right"* because *"if it were to say more, it might have felt like lecturing, and saying less might not have given the message."* P14 suggested leveraging the interactiveness of the question nudge to remind the user to continue asking further questions by proactively offering choices like presenting *"options one through four"* for answering the query. P3 imagined the nudge to guide as a reminder, asking users to *"please check the reference or resource."* Finally, P8 suggested using *"something very strong to let people know this advice is not credible"* in the nudges to signify how the information should be treated, such as the word *"caution."*

4.3.2 Seamless nudges. Many participants (n = 23) gave feedback on the sounds in the non-speech nudges to be non-disruptive and drew a comparison between the serious nature of the conversation with the VA and the lack of urgency exhibited by the non-speech sounds. For example, one of the non-speech nudge prototypes sounded like a video game to P26, and they expressed their confusion by saying, *"It's a sinus infection, man! I'm serious, you know?"* P28 was reminded of a song after listening to one of the sounds and pointed out how *"the other one sounds like a comedic or humorous one for something that is very serious"*. Other participants also echoed similar sentiments, such as how the non-speech nudges sound *"old-sounding, something like what technology would have been 15 years ago"* (P10) or *"silly"* (P7). Additionally, three participants in the migraine scenario raised important concerns about the sounds being unpleasant or *"too loud and brass"* (P12) and how it could potentially make their health condition worse. For example, P33 described how *"playing the music doesn't mean anything, especially if my head is really hurting already."* Interestingly, P14 also pointed to the unfamiliarity of the concept of using sounds to communicate information uncertainty. They elaborated by saying, *"We like big beeps or the airy sound and all that, so we're used to danger. We're not used to a sound being either credible or not credible."*

After being told the purpose of the nudges, three participants reflected on the utility of sounds to communicate information uncertainty. For example, P3 thought that the nudges in the form of non-speech sounds would be very helpful but that *"there would have to be some sort of notification from the company"* relating the tones in the sound to the fact that the information is *"sketchy."* And P29 pointed out that *"I think I prefer the sound better because the nudge becomes very short and simple. And it just gives me the room to proceed with my further inquiries"*. When asked how participants

would improve the sounds, participants stressed the importance of *"something more soothing"* (P28) and *"rhythmic, like the sound that went down the scale"* (P12). This shows that participants appreciated the use of non-speech sounds as long as there was sufficient clarity, and felt that the sounds could be designed more appropriately. As discussed in section 5.1, further research is needed to delineate what such non-speech sounds should sound like.

While the non-speech sounds were designed using characteristics of untrustworthy sounds from the literature, the criticism in these quotes points to the complexity of using sounds to communicate information uncertainty. It shows the need for a robust framework catering to factors such as familiarity, sound characteristics, and context for developing effective sound-based nudges. We use these quotes to further reflect on the tension between the unfamiliarity and utility of using sounds for communicating information uncertainty in the discussion section.

5 Discussion

Older adults and researchers emphasize the need for more critical and iterative search practices with voice technologies, as well as quicker recovery mechanisms to address information uncertainty [16, 18, 76, 122, 123]. However, current VA capabilities fail to provide the necessary dialogue to resolve such uncertainties [64]. While HCI researchers have mainly used nudging techniques to achieve mindful interactions [26] in visual experiences, nudging to create such reflective experiences in voice-based systems remains underexplored. Toward that goal, we designed four audio nudge prototypes based on prior research to understand older adults' perceptions (RQ1) and expectations with audio nudges in supporting critical voice information-seeking practices (RQ2).

Our findings show that speech nudges were more effective at encouraging critical reflection than non-speech nudges because they are more disruptive (RQ1), revealing competing needs between audio as disruptive and as accessible. We then discuss how to design better audio nudges (RQ2), i.e., explanatory and concise nudges, and ensure that they actively promote further interaction and critical dialogue. We contribute to ongoing discussions about addressing information uncertainty for older adults by reflecting on the tensions and design implications for embedding audio-based information quality cues in non-visual interfaces.

5.1 Reflecting on Accessibility and Disruption in Non-visual Interfaces

Prior research on accessibility and audio-based interfaces emphasizes how disruption is detrimental to non-visual information access [24, 39, 59, 75, 99, 115]. For instance, researchers emphasize voice interface design that minimizes cognitive load [13, 68]. Relatedly, researchers propose using non-speech sounds with speech to improve information-seeking experiences. While speech-based interactions facilitate cursory meaning-making (e.g., a natural command-based conversation about hotel booking with a voice-based chatbot) [24], non-speech sounds efficiently enable a quick visual content overview without significant cognitive overload, especially for complex data [25, 44, 96, 97, 100]. Further, research on accessible audio interfaces focuses on intuitiveness, efficiency, and seamlessness [58, 59, 91].

In contrast to this seamlessness, our findings identify disruptiveness as an important characteristic of audio nudges to facilitate critical reflection with older adults. For example, when participants were asked to interpret the purpose of non-speech nudges, they either ignored non-speech sounds or associated them with errors (section 4.2). Comparatively, we observed how participants often paused to discuss information quality when engaging with the question nudge (one of the speech nudges) as the question elicited reflective and cautionary responses. Relatedly, prior research on visual nudges concludes that nudges are effective at prompting critical reflection when they interrupt one's thought process [51, 53]. For example, while warning boxes or labels (e.g., labels warning about potential misinformation beside social posts) are a mild form of disruption in the user's experience, researchers show that they are not sufficient for provoking reflection as they do not address the underlying cognitive biases in the user's decision-making process [31, 53]. Researchers caution against this "magic bullet concept" [56] and encourage designers to focus on mechanisms that elicit critical thinking [22, 26, 53]. Thus, we argue that focusing on disruptiveness remains critical for accessible non-visual nudges.

Our findings show how effective audio nudges prompted older adults to engage in reflective conversations and consider alternate perspectives. While non-speech sounds can be less cognitively demanding than speech sounds [35, 89, 100], our findings show that, once explained, using them in audio nudges required older adults to remember and assign new meanings to sounds. Additionally, recall how some older adult participants reflected that certain sounds could cause discomfort to users with sensory sensitivities or chronic conditions (e.g., migraines) (4.3.2).

As motivated by HCI researchers [18, 79, 83, 100], we argue that non-speech sound design is important beyond meaning-making and critical to accessibility, especially as we move toward a multi-modal information society. In fact, including audio nudges in this multimodal technology landscape raises overarching questions regarding the role of disruptiveness in daily and long-term usage. For instance, while older adults mention the preference for sufficiently disruptive nudges that guide users toward confirming content, the regularity of these audio nudges (consistently repeated across every high-risk query vs. periodic reminders) can affect the experience. This consideration leads to an open question: *What form of regularity is a 'sufficient' disruption?* Similarly, what constitutes 'disruptive' may evolve with the accessibility needs of older adults and habituation with audio nudges (e.g., sound patterns may need to change due to being ignored). For example, Rogers et al. [85] discovered that nudging the use of stairs (instead of the elevator) sustained long-term impact despite its removal, because the reflective engagement significantly disrupted users' routines. Researchers and designers need to consider the equivalent of such disruption for audio nudges by questioning *how 'disruptiveness' relates to long-term attention and usage?*

Thus, considering older adults and people with disabilities are rapidly adopting voice technologies [76, 82, 120], our paper calls for heightened attention on balancing accessibility and encouraging disruptiveness, specifically raising the open question of how to design for disruptive conversation while balancing diverse accessibility needs.

5.2 Designing Reflective Audio Nudges for Older Adults

Although there has been significant research on designing visual cues to communicate information quality and mitigate information uncertainty [53], there is limited to no research on designing accessible audio information quality cues for older adults [6, 19]. Therefore, we also focused on how older adults envision audio nudges to support their information-seeking practices using voice technology. Our findings revealed a consistent need for audio nudges to (1) sufficiently guide older adults to ask critical questions about the VA's response and (2) encourage older adults to confirm information from the VA. We categorize these needs into two distinct yet complementary audio nudge characteristics, *explanatory* and *interactive*. We then expand on how they can make audio nudges more effective in helping older adults mitigate information uncertainty.

Considered as a whole, our findings attempt to answer previous calls from HCI researchers to facilitate critical reflection in voice-based search [10, 16, 18, 76, 111, 123] and advance audio-based nudging [6].

5.2.1 Explanatory Nudges. Researchers have previously shown that older adults want to be active information seekers but find it difficult to confirm or analyze responses without common visual cues (e.g., structural and design elements [104, 121]) [18]. However, our findings illuminate how nudging can potentially lead to undesirable consequences like when the statement nudge ('The information I looked up online related to a health topic') led higher perceived credibility in VA output that contained incorrect information (e.g., suggesting coffee to treat migraine in Table 2). Older adult participants misinterpreted the information source presence as credibility. Prior work has also highlighted possible dangerous outcomes of older adults overtrusting VA responses [10, 28, 111].

We argue that mitigating voice-based information uncertainty through nudging needs to go beyond citing source presence and help users quickly assess information quality. This suggestion is a step towards answering calls in prior work regarding building healthier information ecosystems that encourage and support users in being more critical of information [10, 27, 53], especially older adults [16, 28, 37].

Older adults expected nudges to provide explanation and guidance to help resolve informational uncertainties. So, we draw inspiration from researchers who have used a variety of explanatory and informational nudging to add friction to the user's experience and mitigate the spread of misinformation in visual interfaces. For example, Konstantinou et al. [53] used a statement nudge to indicate that the post about to be shared may contain misinformation and urges users to consider other viewpoints before sharing the post. Similarly, Wang et al. [112] designed a web plugin that aims at mitigating impulsive disclosures on social media by displaying five random users' friends to caution against the post's audience. Researchers argue that such *explanatory* cues help create a momentary interruption that prompts individuals to pause and switch from intuitive, automatic decision-making to a more deliberative and analytical one [26, 53]. However, designing such nudges in audio-based interfaces remains challenging due to the lack of visual affordances or access to screen-based audio interfaces (e.g., being in a different room from the assistant or having low vision).

We believe that audio nudges should **mention key credibility indicators** (e.g., source reputation, ad presence, or fact-check statuses) [48, 118]) researched by HCI scholars to serve as an explanation that signals information quality and sparks interest in further investigation. The audio nudge could compliment this information by suggesting that, *"It is always a good idea to double-check this information with other credible sites"*, to guide users towards alternate viewpoints. Additionally, we note that designers should ensure that these nudges strengthen older adults' agency (i.e., their capacity to deliberately determine what to choose and how to behave) using frameworks posed in prior research such as the "nudges as facilitators, sparks or signals" by Carabal et al [22]. Failing to respect older adults' ability to make rational decisions can hinder their growth as accountable individuals and inhibit critical reflection when searching through voice-based systems [57].

However, doing so brings forth critical open questions regarding the design and implementation of audio nudges: *How can explanatory audio nudges be designed such that they do not cause additional cognitive load for older adults?* And *how can we ensure explanations remain within the conceptual scope of nudges (i.e., differentiate from responses from voice-based systems)*?

5.2.2 Interactive Nudges. The question nudge ('Do you find this believable?') brought forth split perspectives: some older adults appreciated it as a feedback mechanism inviting them to continue the conversation, while others reacted with strong distrust. Upon probing during post-task interviews, this variance interestingly led to a combined view of the question-based framing serving as an effective "spark" (triggers that motivate behavior [33]) or "facilitator" (triggers that make behavior easier [33]) for having critical conversations with the VA. For example, older adults highlighted that the question nudge either encouraged them to look up more information elsewhere or facilitated critical reflection in the form of a follow-up question, ultimately motivating the pursuit of secure information. Researchers argue that *facilitator* and *spark* nudges are effective when a user lacks motivation or knowledge to execute a particular behavior (e.g., resolving uncertainty) [22, 53]. Relatedly, prior research studying older adults' information-searching practices via voice-based systems also suggests similar behavior wherein they are confused about the information received, resulting in gaps in motivation (e.g., avoiding certain types of health queries out of caution) or knowledge (e.g., lacking efficient mechanisms) for mitigating information uncertainty [16, 18, 28].

Thus, given the lack of discoverability facilitated by current voice-based systems despite their popularity amongst older adults [64], we argue that audio nudges should be **designed using a question-based framing** to make older adults more responsive to critical reasoning during audio-based interactions. Researchers focused on utilizing nudging techniques have used a variety of interactive elements to *spark* or *facilitate* desired behavior, such as providing links, using colors, or directing users to the chatbots part of the platform [53]. When designed with intention and introduced with care, such frictions can elicit reflective, informed, and safe interactions. In our context, the explanatory nudge statement suggesting to double-check the information could be combined with a question asking, *"Do you want to get a list of credible websites?"*, as an interactive element that motivates and enables further investigation.

The voice-based system could compliment such suggestions with approaches and metrics in the literature to facilitate further conversation, such as tools that facilitate content assessments through active selection of trusted resources[47]. Such "mindful interactions" [26] can help people transition to mental states where they experience deeper levels of involvement and can guide older adults to have conversations about information quality in voice-based interactions.

However, designing and building such interactive nudges, too, brings forth essential open questions that invite further collaboration between researchers in information retrieval, accessibility, and voice interface design: *Can older adults benefit from personalizing (e.g., framing and length) question-based nudges?* And *how can such nudges be framed to balance overtrust and distrust (a tension that emerged from the disparate reactions to the speech nudges)?*

6 Limitations and Future Work

We acknowledge that there are several limitations to the study. First, we did not tell participants about the purpose of the audio nudges before soliciting feedback to minimize researcher bias and avoid priming. Future studies could consider evaluating prompts to introduce nudges. Doing so can also help acquaint users with unfamiliar nudges (e.g., non-speech audio nudges) and inform appropriate on-boarding language for participants to use audio nudges in their homes.

Second, we focused primarily on reflective nudges due to their non-invasive nature but recognized other nudges may be effective in non-visual forms. Future work could evaluate other forms of sound nudges, such as earcons or hybrid non-speech and speech nudges [46, 71]. Additionally, researchers can also experiment with structured sound design by making characteristics associated with uncertainty more salient to elicit strong reactions. Similarly, we chose one query in a specific context (i.e., health-related information seeking) for each participant to analyze 4 different types of audio nudges. While choosing this high-risk context enabled us to contribute towards one of the common ways older adults use VAs, researchers could study more types of nudges (e.g., differ levels of personalization for speech-based nudges) using multiple queries in different contexts. In doing so, researchers can also vary VA responses for every nudge type, maturing the design space of audio nudges.

Third, while this was a formative study, our recruitment sample consisted of older adults without reported disabilities. Future research can contribute richer design implications that inform accessible design by collecting perceptions from users with visual or cognitive disabilities to collect diverse needs around audio nudging in voice-based interfaces. Fourth, we used pre-recorded answers to mimic a voice assistant and did not facilitate further interaction with the VA prototype. This did not allow participants to ask follow-up questions but did allow for a more controlled environment across participants for credibility assessments. Future work can utilize back-and-forth VA interactions to investigate designing audio nudges for daily (e.g., repetition with queries) and long-term use (e.g., effect of continuous use on attention and disruptiveness).

Lastly, researchers can explore how prior experience with VAs can affect user expectations and perceptions of audio nudges. Understanding these dynamics could help design personalized nudging strategies tailored to users' familiarity and interaction habits with voice-based systems.

7 Conclusion

This paper presented four audio nudge designs and a formative study to evaluate their effectiveness with older adults to help in mitigating information uncertainty via voice-based systems. While we observed that speech-based nudges were more effective than sound-based nudges in prompting critical reflection and encouraging engagement, a deeper analysis of the differences reveals a tension between the need for nudges to be disruptive and accessible. The paper situates this tension in the context of accessibility research on audio-based experiences and discusses design implications for designing effective audio nudges. This work calls researchers in accessibility, conversational design, and nudging to work together to address open questions of mitigating voice-based information uncertainty with VAs.

Acknowledgments

We extend our heartfelt thanks to Rahaf Alharbi, Sam Ankenbauer, Jaylin Herskovitz, Anhong Guo, Jeremy Huang, and Maryam Arab for their invaluable feedback on earlier drafts of this paper. Their insightful comments and thoughtful critiques helped strengthen our analyses and arguments. We also sincerely appreciate Lisa Coyle for her tireless support in supporting our recruitment process. Specially, we thank all the participants for their time and for sharing their perspectives with us. This research would not have been possible without them.

Research reported in this publication was funded by the National Science Foundation under Award # 2143044.

References

[1] Alessandro Acquisti, Idris Adjerid, Rebecca Balebako, Laura Brandimarte, Lorrie Faith Cranor, Saranga Komanduri, Pedro Giovanni Leon, Norman Sadeh, Florian Schaub, Manya Sleeper, et al. 2017. Nudges for privacy and security: Understanding and assisting users' choices online. *ACM Computing Surveys (CSUR)* 50, 3 (2017), 1–41.

[2] Emily Couvillon Alagha and Rachel Renee Helbing. 2019. Evaluating the quality of voice assistants' responses to consumer health questions about vaccines: an exploratory comparison of Alexa, Google Assistant and Siri. *BMJ health & care informatics* 26, 1 (2019), e100075. doi:10.1136/bmjhci-2019-100075

[3] Hugues Ali Mehenni, Sofiya Kobylyanskaya, Ioana Vasilescu, and Laurence Devillers. 2021. Children as Candidates to Verbal Nudging in a Human-robot Experiment. In *Companion Publication of the 2020 International Conference on Multimodal Interaction (ICMI '20 Companion)*. Association for Computing Machinery, New York, NY, USA, 482–486. doi:10.1145/3395035.3425224

[4] Anneliese Arnold, Stephanie Kolody, Aidan Comeau, and Antonio Miguel Cruz. 2022. What does the literature say about the use of personal voice assistants in older adults? A scoping review. *Disability and Rehabilitation: Assistive Technology* 0, 0 (2022), 1–12. doi:10.1080/17483107.2022.2065369 Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/17483107.2022.2065369.

[5] Andrea Ballatore, D. Gordon, and A. P. Boone. 2018. Sonifying data uncertainty with sound dimensions. *Cartography and Geographic Information Science* 46, 5 (2018), 385–400. https://doi.org/10.1080/15230406.2018.1495103 Number: 5 Publisher: Taylor & Francis.

[6] Kristoffer Bergram, Marija Djokovic, Valéry Bezençon, and Adrian Holzer. 2022. The Digital Landscape of Nudging: A Systematic Literature Review of Empirical Research on Digital Nudges. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (2022-04-28) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, 1–16. doi:10.1145/3491102.3517638

[7] Md Momen Bhuiyan, Michael Horning, Sang Won Lee, and Tanushree Mitra. 2021. Nudgecred: Supporting news credibility assessment on social media through nudges. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW2 (2021), 1–30.

[8] Md Momen Bhuiyan, Michael Horning, Sang Won Lee, and Tanushree Mitra. 2021. NudgeCred: Supporting News Credibility Assessment on Social Media Through Nudges. *Proceedings of the ACM on Human-Computer Interaction* 5 (2021), 427:1–427:30. Issue CSCW2. doi:10.1145/3479571

[9] Md Momen Bhuiyan, Kexin Zhang, Kelsey Vick, Michael A. Horning, and Tanushree Mitra. 2018. FeedReflect: A Tool for Nudging Users to Assess News Credibility on Twitter. In *Companion of the 2018 ACM Conference on Computer Supported Cooperative Work and Social Computing* (2018-10-30) *(CSCW '18 Companion)*. Association for Computing Machinery, New York, NY, USA, 205–208. doi:10.1145/3272973.3274056

[10] Timothy W. Bickmore, Ha Trinh, Stefan Olafsson, Teresa K. O'Leary, Reza Asadi, Nathaniel M. Rickles, and Ricardo Cruz. 2018. Patient and Consumer Safety Risks When Using Conversational Assistants for Medical Information: An Observational Study of Siri, Alexa, and Google Assistant. *Journal of Medical Internet Research* 20, 9 (2018), e11510. doi:10.2196/11510 Company: Journal of Medical Internet Research Distributor: Journal of Medical Internet Research Institution: Journal of Medical Internet Research Label: Journal of Medical Internet Research Publisher: JMIR Publications Inc., Toronto, Canada.

[11] Meera M. Blattner, Denise A. Sumikawa, and Robert M. Greenberg. 1989. Earcons and Icons: Their Structure and Common Design Principles. *Human–Computer Interaction* 4, 1 (March 1989), 11–44. doi:10.1207/s15327051hci0401_1

[12] Karen Bonilla, Brian Gaitan, Jamie Sanders, Noami Khenglawt, and Aqueasha Martin-Hammond. 2021. Comparing Older and Younger Adults Perceptions of Voice and Text-based Search for Consumer Health Information Tasks. *AMIA Annual Symposium Proceedings* 2021 (2021), 227–236. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8861708/

[13] Stacy M. Branham and Antony Rishin Mukkath Roy. 2019. Reading Between the Guidelines: How Commercial Voice Assistant Guidelines Hinder Accessibility for Blind Users. In *Proceedings of the 21st International ACM SIGACCESS Conference on Computers and Accessibility* (2019-10-24) *(ASSETS '19)*. Association for Computing Machinery, New York, NY, USA, 446–458. doi:10.1145/3308561.3353797

[14] Nadia M. Brashier and Daniel L. Schacter. 2020. Aging in an Era of Fake News. *Current Directions in Psychological Science* 29, 3 (2020), 316–323. doi:10.1177/0963721420915872 Publisher: SAGE Publications Inc.

[15] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology* 3, 2 (2006), 77–101. doi:10.1191/1478088706qp063oa Publisher: Routledge _eprint: https://www.tandfonline.com/doi/pdf/10.1191/1478088706qp063oa.

[16] Robin Brewer. 2023. Understanding voice-based information uncertainty: A case study of health information seeking with voice assistants. *Journal of the Association for Information Science and Technology* n/a (2023). Issue n/a. doi:10.1002/asi.24854 _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/asi.24854.

[17] Robin Brewer, Raymundo Cornejo Garcia, Tedmond Schwaba, Darren Gergle, and Anne Marie Piper. 2016. Exploring Traditional Phones as an E-Mail Interface for Older Adults. *ACM Trans. Access. Comput.* 8, 2 (Jan. 2016), 6:1–6:20. doi:10.1145/2839303

[18] Robin Brewer, Casey Pierce, Pooja Upadhyay, and Leeseul Park. 2022. An Empirical Study of Older Adult's Voice Assistant Use for Health Information Seeking. *ACM Transactions on Interactive Intelligent Systems* 12, 2 (2022), 13:1–13:32. doi:10.1145/3484507

[19] L. M. Brown, S. A. Brewster, S. A. Ramloll, R. Burton, and B. Riedel. 2003. Design Guidelines for Audio Presentation of Graphs and Tables. http://eprints.gla.ac.uk/3196/, 284–287 pages.

[20] Romain Cadario and Pierre Chandon. 2020. Which healthy eating nudges work best? A meta-analysis of field experiments. *Marketing Science* 39, 3 (2020), 465–486.

[21] Clara Caldeira, Matthew Bietz, Marisol Vidauri, and Yunan Chen. 2017. Senior Care for Aging in Place: Balancing Assistance and Independence. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing (CSCW '17)*. Association for Computing Machinery, New York, NY, USA, 1605–1617. doi:10.1145/2998181.2998206

[22] Ana Caraban, Evangelos Karapanos, Daniel Gonçalves, and Pedro Campos. 2019. 23 Ways to Nudge: A Review of Technology-Mediated Nudging in Human-Computer Interaction. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (2019-05-02) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–15. doi:10.1145/3290605.3300733

[23] Eun Cheol Choi and Emilio Ferrara. 2024. Automated claim matching with large language models: empowering fact-checkers in the fight against misinformation. In *Companion Proceedings of the ACM on Web Conference 2024*. 1441–1449.

[24] Leigh Clark, Philip Doyle, Diego Garaialde, Emer Gilmartin, Stephan Schlögl, Jens Edlund, Matthew Aylett, João Cabral, Cosmin Munteanu, Justin Edwards, and Benjamin R Cowan. 2019. The State of Speech in HCI: Trends, Themes and Challenges. *Interacting with Computers* 31, 4 (2019), 349–371. doi:10.1093/iwc/iwz016

[25] Matthew Clark and Afsaneh Doryab. 2023. Sounds of Health: Using Personalized Sonification Models to Communicate Health Information. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 6, 4 (2023), 206:1–206:31. doi:10.1145/3570346

[26] Anna L. Cox, Sandy J.J. Gould, Marta E. Cecchinato, Ioanna Iacovides, and Ian Renfree. 2016. Design Frictions for Mindful Interactions: The Case for Microboundaries. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '16)*. Association for Computing Machinery, New York, NY, USA, 1389–1397. doi:10.1145/2851581.2892410

[27] Henry Kudzanai Dambanemuya and Nicholas Diakopoulos. 2021. Auditing the Information Quality of News-Related Queries on the Alexa Voice Assistant. *Proceedings of the ACM on Human-Computer Interaction* 5 (2021), 83:1–83:21. Issue CSCW1. doi:10.1145/3449157

[28] Smit Desai and Jessie Chin. 2023. OK Google, Let's Learn: Using Voice User Interfaces for Informal Self-Regulated Learning of Health Topics among Younger and Older Adults. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23)*. Association for Computing Machinery, New York, NY, USA, 1–21. doi:10.1145/3544548.3581507

[29] Alan Dix. 2020. *Statistics for HCI: making sense of quantitative data*. Morgan & Claypool Publishers.

[30] Spencer Early, Seyedmohammadmahdi Mirhoseini, Nour El Shamy, and Khaled Hassanein. 2020. Relying on System 1 Thinking Leaves You Susceptible to the Peril of Misinformation. In *Information Systems and Neuroscience*, Fred D. Davis, René Riedl, Jan vom Brocke, Pierre-Majorique Léger, Adriane B. Randolph, and Thomas Fischer (Eds.). Springer International Publishing, Cham, 42–48. doi:10.1007/978-3-030-60073-0_5

[31] Miriam Fernandez and Harith Alani. 2018. Online Misinformation: Challenges and Future Directions. In *Companion of the The Web Conference 2018 on The Web Conference 2018 - WWW '18*. ACM Press, Lyon, France, 595–602. doi:10.1145/3184558.3188730

[32] Andrew J Flanagin and Miriam J Metzger. 2007. The role of site features, user attributes, and information verification behaviors on the perceived credibility of web-based information. *New media & society* 9, 2 (2007), 319–342.

[33] BJ Fogg. 2009. A Behavior Model for Persuasive Design. In *Proceedings of the 4th International Conference on Persuasive Technology (Persuasive '09)*. Association for Computing Machinery, New York, NY, USA, 1–7. doi:10.1145/1541948.1541999

[34] Norman M Fraser and G Nigel Gilbert. 1991. Simulating speech systems. *Computer Speech & Language* 5, 1 (1991), 81–99.

[35] Euan Freeman, Graham Wilson, Dong-Bach Vo, Alex Ng, Ioannis Politis, and Stephen Brewster. 2017. Multimodal feedback in HCI: haptics, non-speech audio, and their applications. In *The Handbook of Multimodal-Multisensor Interfaces: Foundations, User Modeling, and Common Modality Combinations - Volume 1*. Vol. 14. Association for Computing Machinery and Morgan & Claypool, 277–317. https://doi.org/10.1145/3015783.3015792

[36] Stavros Garzonis, Simon Jones, Tim Jay, and Eamonn O'Neill. 2009. Auditory Icon and Earcon Mobile Service Notifications: Intuitiveness, Learnability, Memorability and Preference. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '09)*. Association for Computing Machinery, New York, NY, USA, 1513–1522. doi:10.1145/1518701.1518932

[37] Amira Ghenai., Xueguang Ma., Robin Cohen., Karyn Moffatt., Andy Yang., and Yipeng Ji. 2021. e-Health for Older Adults: Navigating Misinformation. In *Proceedings of the 7th International Conference on Information and Communication Technologies for Ageing Well and e-Health - ICT4AWE*. INSTICC, SciTePress, 195–202. doi:10.5220/0010463001950202

[38] Marcel Gohsen, Johannes Kiesel, Mariam Korashi, Jan Ehlers, and Benno Stein. 2023. Guiding Oral Conversations: How to Nudge Users Towards Asking Questions?. In *Proceedings of the 2023 Conference on Human Information Interaction and Retrieval* (2023-03-20) *(CHIIR '23)*. Association for Computing Machinery, New York, NY, USA, 34–42. doi:10.1145/3576840.3578291

[39] Gabriel Haas, Michael Rietzler, Matt Jones, and Enrico Rukzio. 2022. Keep it Short: A Comparison of Voice Assistants' Response Behavior. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (2022-04-29) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, 1–12. doi:10.1145/3491102.3517684

[40] Pelle Guldborg Hansen and Andreas Jespersen. 2013. Nudge and the Manipulation of Choice. A Framework for the Responsible Use of Nudge Approach to Behaviour Change in Public Policy. https://papers.ssrn.com/abstract=2555337

[41] Christina N Harrington and Lisa Egede. 2023. Trust, comfort and relatability: Understanding black older adults' perceptions of chatbot design for health information seeking. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–18.

[42] Tianzhi He, Farrokh Jazizadeh, and Laura Arpan. 2022. AI-powered virtual assistants nudging occupants for energy saving: proactive smart speakers for HVAC control. *Building Research & Information* 50, 4 (2022), 394–409. doi:10.1080/09613218.2021.2012119 Publisher: Routledge _eprint: https://doi.org/10.1080/09613218.2021.2012119.

[43] Thomas Hermann, Andy Hunt, and John Neuhoff. 2011. *The Sonification Handbook*. Logos Verlag Berlin.

[44] Leona M Holloway, Cagatay Goncu, Alon Ilsar, Matthew Butler, and Kim Marriott. 2022. Infosonics: Accessible Infographics for People who are Blind using Sonification and Voice. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (2022-04-29) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, 1–13. doi:10.1145/3491102.3517465

[45] Benjamin D. Horne, Mauricio Gruppi, and Sibel Adali. 2019. Trustworthy Misinformation Mitigation with Soft Information Nudging. In *2019 First IEEE International Conference on Trust, Privacy and Security in Intelligent Systems and Applications (TPS-ISA)* (2019-12). IEEE Computer Society, Los Alamitos, CA, USA, 245–254. doi:10.1109/TPS-ISA48467.2019.00039

[46] Ibrar Hussain, Ling Chen, Hamid Turab Mirza, Gencai Chen, and Saeed-Ul Hassan. 2015. Right mix of speech and non-speech: hybrid auditory feedback in mobility assistance of the visually impaired. *Univers. Access Inf. Soc.* 14, 4 (nov 2015), 527–536. doi:10.1007/s10209-014-0350-7

[47] Farnaz Jahanbakhsh and David R Karger. 2024. A Browser Extension for In-Place Signaling and Assessment of Misinformation. In *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '24)*. Association for Computing Machinery, New York, NY, USA, 1–21. doi:10.1145/3613904.3642473

[48] Farnaz Jahanbakhsh, Amy X. Zhang, and David R. Karger. 2022. Leveraging Structured Trusted-Peer Assessments to Combat Misinformation. *Proc. ACM Hum.-Comput. Interact.* 6, CSCW2 (Nov. 2022), 524:1–524:40. doi:10.1145/3555637

[49] Myounghoon Jeon and Bruce N. Walker. 2011. Spindex (Speech Index) Improves Auditory Menu Acceptance and Navigation Performance. *ACM Transactions on Accessible Computing* 3, 3 (2011), 10:1–10:26. doi:10.1145/1952383.1952385

[50] Kaixin Ji, Sachin Pathiyan Cherumanal, Johanne R. Trippas, Danula Hettiachchi, Flora D. Salim, Falk Scholer, and Damiano Spina. 2024. Towards Detecting and Mitigating Cognitive Bias in Spoken Conversational Search. https://arxiv.org/abs/2405.12480v2. doi:10.1145/3640471.3680245

[51] Jan Keller, Tobias Herrmann-Schwarz, Christina Roitzheim, Lea Mertens, Lina Christin Brockmeier, and Aditya Kumar Purohit. 2024. A Digital Nudge-Based Intervention to Interrupt Instagram Usage. *European Journal of Health Psychology* (Aug. 2024). doi:10.1027/2512-8442/a000150

[52] Sunyoung Kim and Abhishek Choudhury. 2021. Exploring older adults' perception and use of smart speaker-based voice assistants: A longitudinal study. *Computers in Human Behavior* 124 (2021). Issue C. doi:10.1016/j.chb.2021.106914

[53] Loukas Konstantinou, Dionysis Panos, and Evangelos Karapanos. 2024. Exploring the Design of Technology-Mediated Nudges for Online Misinformation. *International Journal of Human–Computer Interaction* 0, 0 (2024), 1–28. doi:10.1080/10447318.2023.2301265 Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/10447318.2023.2301265.

[54] KP Krishna Kumar and G Geethakumari. 2014. Detecting misinformation in online social networks using cognitive psychology. *Human-centric Computing and Information Sciences* 4, 1 (2014), 14.

[55] David MJ Lazer, Matthew A Baum, Yochai Benkler, Adam J Berinsky, Kelly M Greenhill, Filippo Menczer, Miriam J Metzger, Brendan Nyhan, Gordon Pennycook, David Rothschild, et al. 2018. The science of fake news. *Science* 359, 6380 (2018), 1094–1096.

[56] Kalev Leetaru. 2017. *The Backfire Effect And Why Facebook's 'Fake News' Warning Gets It All Wrong*. https://www.forbes.com/sites/kalevleetaru/2017/03/23/the-backfire-effect-and-why-facebooks-fake-news-warning-gets-it-all-wrong/ Section: Tech.

[57] Neil Levy. 2017. Nudges in a post-truth world. *Journal of Medical Ethics* 43, 8 (2017), 495–500. doi:10.1136/medethics-2017-104153 Publisher: Institute of Medical Ethics Section: Extended essay.

[58] Kate Lister, Tim Coughlan, Francisco Iniesto, Nick Freear, and Peter Devine. 2020. Accessible Conversational User Interfaces: Considerations for Design. In *Proceedings of the 17th International Web for All Conference (W4A '20)*. Association for Computing Machinery, New York, NY, USA, 1–11. doi:10.1145/3371300.3383343

[59] Tao Long and Lydia B. Chilton. 2023. Challenges and Opportunities for the Design of Smart Speakers. doi:10.48550/arXiv.2306.05741 arXiv:2306.05741 [cs]

[60] Andrei Lăpușteanu, Anca Morar, Alin Moldoveanu, Maria-Anca Bălușoiu, and Florica Moldoveanu. 2024. A review of sonification solutions in assistive systems for visually impaired people. *Disability and Rehabilitation: Assistive Technology* 0, 0 (2024), 1–16. doi:10.1080/17483107.2024.2326590 Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/17483107.2024.2326590.

[61] Douglas MacFarlane, Mark J. Hurlstone, and Ullrich K. H. Ecker. 2020. Protecting consumers from fraudulent health claims: A taxonomy of psychological drivers, interventions, barriers, and treatments. *Social Science & Medicine (1982)* 259 (2020), 112790. doi:10.1016/j.socscimed.2020.112790

[62] Hiroaki Masaki, Kengo Shibata, Shui Hoshino, Takahiro Ishihama, Nagayuki Saito, and Koji Yatani. 2020. Exploring nudge designs to help adolescent SNS users avoid privacy and safety threats. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–11.

[63] Jeffrey Maynes. 2015. Critical Thinking and Cognitive Bias. *Informal Logic* 35, 2 (May 2015), 183–203. doi:10.22329/il.v35i2.4187

[64] Indra Mckie, Bhuva Narayan, and Baki Kocaballi. 2022. Conversational Voice Assistants and a Case Study of Long-Term Users: A Human Information Behaviours Perspective. *Journal of the Australian Library and Information Association* 71, 3 (July 2022), 233–255. doi:10.1080/24750158.2022.2104738

[65] Maria D. Molina, S. Shyam Sundar, Thai Le, and Dongwon Lee. 2021. "Fake News" Is Not Simply False Information: A Concept Explication and Taxonomy of Online Content. *American Behavioral Scientist* 65, 2 (2021), 180–212. doi:10.1177/0002764219878224 Publisher: SAGE Publications Inc.

[66] Justin C. Morse, Ping Li, Kim A. Ely, Meghan H. Shilts, Todd J. Wannemuehler, Li-Ching Huang, Quanhu Sheng, Naweed I. Chowdhury, Rakesh K. Chandra, Suman R. Das, and Justin H. Turner. 2019. Chronic rhinosinusitis in elderly patients is associated with an exaggerated neutrophilic proinflammatory response to pathogenic bacteria. *Journal of Allergy and Clinical Immunology* 143, 3 (2019), 990–1002.e6. doi:10.1016/j.jaci.2018.10.056

[67] Emadoddin Moudi, Seyed Reza Hosseini, and Ali Bijani. 2017. Nephrolithiasis in elderly population; effect of demographic characteristics. *Journal of Nephropathology* 6, 2 (2017), 63–68. doi:10.15171/jnp.2017.11

[68] Christine Murad, Cosmin Munteanu, Leigh Clark, and Benjamin R. Cowan. 2018. Design guidelines for hands-free speech interaction. In *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct* (2018-09-03) *(MobileHCI '18)*. Association for Computing Machinery, New York, NY, USA, 269–276. doi:10.1145/3236112.3236149

[69] Emil Myresten, David Larson Holmgren, and Roberto Bresin. 2021. Sonification of Twitter Hashtags Using Earcons Based on the Sound of Vowels. (2021). doi:10.5281/zenodo.5723812 Publisher: Zenodo.

[70] Preslav Nakov, David Corney, Maram Hasanain, Firoj Alam, Tamer Elsayed, Alberto Barrón-Cedeño, Paolo Papotti, Shaden Shaar, and Giovanni Da San Martino. 2021. Automated fact-checking for assisting human fact-checkers. *arXiv preprint arXiv:2103.07769* (2021).

[71] Michael A. Nees and Eliana Liebman. 2023. Auditory Icons, Earcons, Spearcons, and Speech: A Systematic Review and Meta-Analysis of Brief Audio Alerts in Human-Machine Interfaces. *Auditory Perception & Cognition* 6, 3 (2023), 300–329. doi:10.1080/25742442.2023.2219201 Publisher: Routledge _eprint: https://doi.org/10.1080/25742442.2023.2219201.

[72] Elmie Nekmat. 2020. Nudge Effect of Fact-Check Alerts: Source Influence and Media Skepticism on Sharing of News Misinformation in Social Media. *Social Media + Society* 6, 1 (2020), 2056305119897322. doi:10.1177/2056305119897322 Publisher: SAGE Publications Ltd.

[73] Katherine M. Nelson, Mirja Kristina Bauer, and Stefan Partelow. 2021. Informational Nudges to Encourage Pro-environmental Behavior: Examining Differences in Message Framing and Human Interaction. *Frontiers in Communication* 5 (Feb. 2021). doi:10.3389/fcomm.2020.610186

[74] Novia Nurain, Chia-Fang Chung, Clara Caldeira, and Kay Connelly. 2021. Hugging with a Shower Curtain: Older Adults' Social Support Realities During the COVID-19 Pandemic. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW2 (2021), 1–31.

[75] Cathy Pearl. 2016. *Designing Voice User Interfaces: Principles of Conversational Experiences*. "O'Reilly Media, Inc.". Google-Books-ID: MmnEDQAAQBAJ.

[76] Sayli Pednekar, Palka Dhirawani, Richa Shah, Narendra Shekokar, and Kranti Ghag. 2023. Voice-Based Interaction for an Aging Population: A Systematic Review. In *2023 3rd International Conference on Intelligent Communication and Computational Techniques (ICCT)* (2023-01). 1–8. doi:10.1109/ICCT56969.2023.10075801

[77] Gordon Pennycook, Jonathon McPhetres, Yunhao Zhang, Jackson G. Lu, and David G. Rand. 2020. Fighting COVID-19 Misinformation on Social Media: Experimental Evidence for a Scalable Accuracy-Nudge Intervention. *Psychological Science* 31, 7 (2020), 770–780. doi:10.1177/0956797620939054 Publisher: SAGE Publications Inc.

[78] Kashyap Popat, Subhabrata Mukherjee, Jannik Strötgen, and Gerhard Weikum. 2017. Where the truth lies: Explaining the credibility of emerging claims on the web and social media. In *Proceedings of the 26th international conference on world wide web companion*. 1003–1012.

[79] Venkatesh Potluri, Maulishree Pandey, Andrew Begel, Michael Barnett, and Scott Reitherman. 2022. In CodeWalk: Facilitating Shared Awareness in Mixed-Ability Collaborative Software Development (2022-10-22). *Proceedings of the 24th International ACM SIGACCESS Conference on Computers and Accessibility* 20, 1–16. doi:10.1145/3517428.3544812

[80] Alisha Pradhan, Leah Findlater, and Amanda Lazar. 2019. "Phantom Friend" or "Just a Box with Information": Personification and Ontological Categorization of Smart Speaker-based Voice Assistants by Older Adults. *Proceedings of the ACM on Human-Computer Interaction* 3 (2019), 214:1–214:21. Issue CSCW. doi:10.1145/3359316

[81] Alisha Pradhan, Amanda Lazar, and Leah Findlater. 2020. Use of Intelligent Voice Assistants by Older Adults with Low Technology Use. *ACM Transactions on Computer-Human Interaction* 27, 4 (2020), 31:1–31:27. doi:10.1145/3373759

[82] Alisha Pradhan, Kanika Mehta, and Leah Findlater. 2018. " Accessibility Came by Accident" Use of Voice-Controlled Intelligent Personal Assistants by People with Disabilities. In *Proceedings of the 2018 CHI Conference on human factors in*

computing systems. 1–13.

[83] Lauren Race, Kia El-Amin, Sarah Anoke, Andrew Hayward, Amber James, Amy Hurst, Audrey Davis, and Theresa Mershon. 2022. Understanding Design Preferences for Sensory-Sensitive Earcons with Neurodivergent Individuals. In *Proceedings of the 24th International ACM SIGACCESS Conference on Computers and Accessibility* (2022-10-22) *(ASSETS '22)*. Association for Computing Machinery, New York, NY, USA, 1–4. doi:10.1145/3517428.3550365

[84] Nina Riggins and Annika Ehrlich. 2022. Episodic Migraine and Older Adults. *Current Pain and Headache Reports* 26, 4 (2022), 331–335. doi:10.1007/s11916-022-01029-7

[85] Yvonne Rogers, William R. Hazlewood, Paul Marshall, Nick Dalton, and Susanna Hertrich. 2010. Ambient influence: can twinkly lights lure and abstract representations trigger behavioral change?. In *Proceedings of the 12th ACM International Conference on Ubiquitous Computing* (Copenhagen, Denmark) *(UbiComp '10)*. Association for Computing Machinery, New York, NY, USA, 261–270. doi:10.1145/1864349.1864372

[86] William Samuelson and Richard Zeckhauser. 1988. Status Quo Bias in Decision Making. *Journal of Risk and Uncertainty* 1, 1 (1988), 7–59. jstor:41760530

[87] Laura D. Scherer and Gordon Pennycook. 2020. Who Is Susceptible to Online Health Misinformation? *American Journal of Public Health* 110 (2020), S276–S277. Issue S3. doi:10.2105/AJPH.2020.305908 Publisher: American Public Health Association.

[88] Matthias Schmidmaier, Heinrich Hußmann, and Dominik Maurice Runge. 2020. Beep Beep: Building Trust with Sound. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems* (2020-04-25) *(CHI EA '20)*. Association for Computing Machinery, New York, NY, USA, 1–8. doi:10.1145/3334480.3382848

[89] Björn W. Schuller, Tuomas Virtanen, Maria Riveiro, Georgios Rizos, Jing Han, Annamaria Mesaros, and Konstantinos Drossos. 2021. Towards Sonification in Multimodal and User-friendlyExplainable Artificial Intelligence. In *Proceedings of the 2021 International Conference on Multimodal Interaction* (2021-10-18). ACM, Montréal QC Canada, 788–792. doi:10.1145/3462244.3479879

[90] Marlene B Schwartz. 2007. The influence of a verbal prompt on school lunch fruit consumption: a pilot study. *International Journal of Behavioral Nutrition and Physical Activity* 4 (2007), 1–5.

[91] Eliseo Sciarretta and Lia Alimenti. 2021. Smart Speakers for Inclusion: How Can Intelligent Virtual Assistants Really Assist Everybody?. In *Human-Computer Interaction. Theory, Methods and Tools*, Masaaki Kurosu (Ed.). Springer International Publishing, Cham, 77–93. doi:10.1007/978-3-030-78462-1_6

[92] JooYoung Seo and Megan Rogge. 2023. Coding Non-Visually in Visual Studio Code: Collaboration Towards Accessible Development Environment for Blind Programmers. In *Proceedings of the 25th International ACM SIGACCESS Conference on Computers and Accessibility* (2023-10-22) *(ASSETS '23)*. Association for Computing Machinery, New York, NY, USA, 1–9. doi:10.1145/3597638.3614550

[93] William Seymour and Max Van Kleek. 2021. Exploring Interactions Between Trust, Anthropomorphism, and Relationship Development in Voice Assistants. *Proceedings of the ACM on Human-Computer Interaction* 5 (2021), 371:1–371:16. Issue CSCW2. doi:10.1145/3479515

[94] Emre Sezgin, Yungui Huang, Ujjwal Ramtekkar, and Simon Lin. 2020. Readiness for voice assistants to support healthcare delivery during a health crisis and pandemic. *npj Digital Medicine* 3, 1 (2020), 1–4. doi:10.1038/s41746-020-00332-0 Publisher: Nature Publishing Group.

[95] Esha Shandilya and Mingming Fan. 2022. Understanding older adults' perceptions and challenges in using AI-enabled everyday technologies. In *Proceedings of the Tenth International Symposium of Chinese CHI*. 105–116.

[96] Ather Sharif, Olivia H. Wang, and Alida T. Muongchan. 2022. "What Makes Sonification User-Friendly?" Exploring Usability and User-Friendliness of Sonified Responses. In *Proceedings of the 24th International ACM SIGACCESS Conference on Computers and Accessibility* (2022-10-22) *(ASSETS '22)*. Association for Computing Machinery, New York, NY, USA, 1–5. doi:10.1145/3517428.3550360

[97] Ather Sharif, Olivia H. Wang, Alida T. Muongchan, Katharina Reinecke, and Jacob O. Wobbrock. 2022. VoxLens: Making Online Data Visualizations Accessible with an Interactive JavaScript Plug-In. In *CHI Conference on Human Factors in Computing Systems* (2022-04-29). ACM, New Orleans LA USA, 1–19. doi:10.1145/3491102.3517431

[98] Ather Sharif, Ruican Zhong, and Yadi Wang. 2023. Conveying Uncertainty in Data Visualizations to Screen-Reader Users Through Non-Visual Means. In *Proceedings of the 25th International ACM SIGACCESS Conference on Computers and Accessibility* (2023-10-22) *(ASSETS '23)*. Association for Computing Machinery, New York, NY, USA, 1–6. doi:10.1145/3597638.3614502

[99] Lei Shi, Brianna J. Tomlinson, John Tang, Edward Cutrell, Daniel McDuff, Gina Venolia, Paul Johns, and Kael Rowan. 2019. Accessible Video Calling: Enabling Nonvisual Perception of Visual Conversation Cues. *Proceedings of the ACM on Human-Computer Interaction* 3 (2019), 131:1–131:22. Issue CSCW. doi:10.1145/3359233

[100] Alexa Siu, Gene S-H Kim, Sile O'Modhrain, and Sean Follmer. 2022. Supporting Accessible Data Visualization Through Audio Data Narratives. In *CHI Conference on Human Factors in Computing Systems* (2022-04-29). ACM, New Orleans LA USA, 1–19. doi:10.1145/3491102.3517678

[101] JANAN AL-AWAR SMITHER. 1993. Short term memory demands in processing synthetic speech by old and young adults. *Behaviour & Information Technology* 12, 6 (1993), 330–335. doi:10.1080/01449299308924397 Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/01449299308924397.

[102] Brendan Spillane, Emer Gilmartin, Christian Saam, and Vincent Wade. 2019. Issues relating to trust in care agents for the elderly. In *Proceedings of the 1st International Conference on Conversational User Interfaces* (Dublin, Ireland) *(CUI '19)*. Association for Computing Machinery, New York, NY, USA, Article 20, 3 pages. doi:10.1145/3342775.3342808

[103] Besiki Stvilia, Lorri Mon, and Yong Jeong Yi. 2009. A model for online consumer health information quality. *Journal of the American Society for Information Science and Technology* 60, 9 (2009), 1781–1791. doi:10.1002/asi.21115 _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/asi.21115.

[104] S. Shyam Sundar, Silvia Knobloch-Westerwick, and Matthias R. Hastall. 2007. News Cues: Information Scent and Cognitive Heuristics. *Journal of the American Society for Information Science and Technology* 58, 3 (2007), 366–378. doi:10.1002/asi.20511

[105] Cass R Sunstein. 2015. Nudging and choice architecture: Ethical considerations. *Yale Journal on Regulation, Forthcoming* (2015).

[106] Siu-Lan Tan, Annabel Cohen, Annabel J. Cohen, Scott D. Lipscomb, and Roger A. Kendall. 2013. *The Psychology of Music in Multimedia*. OUP Oxford. Google-Books-ID: 89m5EyX59ugC.

[107] Richard H. Thaler and Cass R. Sunstein. 2009. *Nudge: Improving Decisions About Health, Wealth, and Happiness*. Penguin. Google-Books-ID: NGA9DwAAQBAJ.

[108] Ratan Bahadur Thapa, Mexhid Ferati, and G. Anthony Giannoumis. 2017. Using non-speech sounds to increase web image accessibility for screen-reader users. In *Proceedings of the 35th ACM International Conference on the Design of Communication* (2017-08-11) *(SIGDOC '17)*. Association for Computing Machinery, New York, NY, USA, 1–9. doi:10.1145/3121113.3121231

[109] Calum Thornhill, Quentin Meeus, Jeroen Peperkamp, and Bettina Berendt. 2019. A digital nudge to counter confirmation bias. *Frontiers in big data* 2 (2019), 11.

[110] Kenneth E Train. 2009. *Discrete choice methods with simulation*. Cambridge university press.

[111] Pooja Upadhyay, Sharon Heung, Shiri Azenkot, and Robin N. Brewer. 2023. Studying Exploration & Long-Term Use of Voice Assistants by Older Adults. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (2023-04-19) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA, 1–11. doi:10.1145/3544548.3580925

[112] Yang Wang, Pedro Giovanni Leon, Alessandro Acquisti, Lorrie Faith Cranor, Alain Forget, and Norman Sadeh. 2014. A Field Trial of Privacy Nudges for Facebook. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*. Association for Computing Machinery, New York, NY, USA, 2367–2376. doi:10.1145/2556288.2557413

[113] Yang Wang, Pedro Giovanni Leon, Kevin Scott, Xiaoxuan Chen, Alessandro Acquisti, and Lorrie Faith Cranor. 2013. Privacy nudges for social media: an exploratory Facebook study. In *Proceedings of the 22nd international conference on world wide web*. 763–770.

[114] Justin D Weisz, Jessica He, Michael Muller, Gabriela Hoefer, Rachel Miles, and Werner Geyer. 2024. Design Principles for Generative AI Applications. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–22.

[115] Linda Wulf, Markus Garschall, Julia Himmelsbach, and Manfred Tscheligi. 2014. Hands free - care free: elderly people taking advantage of speech-only interaction. In *Proceedings of the 8th Nordic Conference on Human-Computer Interaction: Fun, Fast, Foundational* (2014-10-26) *(NordiCHI '14)*. Association for Computing Machinery, New York, NY, USA, 203–206. doi:10.1145/2639189.2639251

[116] Honglian Xiang, Jia Zhou, and Zhuowen Wang. 2023. Reducing Younger and Older Adults' Engagement with COVID-19 Misinformation: The Effects of Accuracy Nudge and Exogenous Cues. *International Journal of Human–Computer Interaction* 0, 0 (2023), 1–16. doi:10.1080/10447318.2022.2158263 Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/10447318.2022.2158263.

[117] Xiao Zhan, Noura Abdi, William Seymour, and Jose Such. 2024. Healthcare Voice AI Assistants: Factors Influencing Trust and Intention to Use. doi:10.1145/3637339 arXiv:2401.04543 [cs]

[118] Amy X. Zhang, Aditya Ranganathan, Sarah Emlen Metz, Scott Appling, Connie Moon Sehat, Norman Gilmore, Nick B. Adams, Emmanuel Vincent, Jennifer Lee, Martin Robbins, Ed Bice, Sandro Hawke, David Karger, and An Xiao Mina. 2018. A Structured Response to Misinformation: Defining and Annotating Credibility Indicators in News Articles. In *Companion Proceedings of the The Web Conference 2018 (WWW '18)*. International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, CHE, 603–612. doi:10.1145/3184558.3188731

[119] Yichi Zhang, Edmund W. J. Lee, and Wei-Peng Teo. 2023. Health-Seeking Behavior and Its Associated Technology Use: Interview Study Among Community-Dwelling Older Adults. *JMIR Aging* 6, 1 (2023), e43709. doi:10.2196/43709 Company: JMIR Aging Distributor: JMIR Aging Institution: JMIR Aging Label: JMIR Aging Publisher: JMIR Publications Inc., Toronto, Canada.

[120] Yan Zhang, Jiaying Liu, and Shijie Song. 2023. The design and evaluation of a nudge-based interface to facilitate consumers' evaluation of online health information credibility. *Journal of the Association for Information Science and Technology* 74, 7 (2023), 828–845. doi:10.1002/asi.24759 _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/asi.24759.

[121] Yan Zhang, Yalin Sun, and Bo Xie. 2015. Quality of health information for consumers on the web: A systematic review of indicators, criteria, tools, and evaluation results. *Journal of the Association for Information Science and Technology* 66, 10 (2015), 2071–2084. doi:10.1002/asi.23311 _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/asi.23311.

[122] Yuxiang Chris Zhao, Mengyuan Zhao, and Shijie Song. 2022. Online Health Information Seeking Behaviors Among Older Adults: Systematic Scoping Review. *Journal of Medical Internet Research* 24, 2 (2022), e34790. doi:10.2196/34790 Company: Journal of Medical Internet Research Distributor: Journal of Medical Internet Research Institution: Journal of Medical Internet Research Label: Journal of Medical Internet Research Publisher: JMIR Publications Inc., Toronto, Canada.

[123] Jia Zhou, Honglian Xiang, and Bingjun Xie. 2023. Better safe than sorry: a study on older adults' credibility judgments and spreading of health misinformation. *Universal Access in the Information Society* 22, 3 (2023), 957–966. doi:10.1007/s10209-022-00899-3